



The Influence of Roughness on Gear Surface Fatigue

Timothy L. Krantz

U.S. Army Research Laboratory, Glenn Research Center, Cleveland, Ohio

The NASA STI Program Office . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the Lead Center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at <http://www.sti.nasa.gov>
- E-mail your question via the Internet to help@sti.nasa.gov
- Fax your question to the NASA Access Help Desk at 301-621-0134
- Telephone the NASA Access Help Desk at 301-621-0390
- Write to:
NASA Access Help Desk
NASA Center for Aerospace Information
7121 Standard Drive
Hanover, MD 21076



The Influence of Roughness on Gear Surface Fatigue

Timothy L. Krantz

U.S. Army Research Laboratory, Glenn Research Center, Cleveland, Ohio

National Aeronautics and
Space Administration

Glenn Research Center

Acknowledgments

I would first especially like to thank my dissertation advisor, Thomas P. Kicher, for his encouragement, guidance and advice. I also thank the members of the reviewing committee: Maurice Adams, Clare Rimnac, and Robert Mullen. I thank my employer, the Army Research Laboratory, Vehicle Technology Directorate, NASA Glenn Research Center, for placing high value on employee education and for supporting this work. Also, this work was possible only with the support of NASA Glenn Research Center staff, especially the Mechanical Components Branch. I thank my supervisors during the duration of this project for their encouragement and guidance: Robert Bill, James Zakrajsek, George Bobula, and John Coy. Special thanks go to Dennis Townsend for permitting the use of specimens from his stock of test gears and for his advice concerning gear fatigue testing. The extensive and high quality of gear fatigue research he conducted during his career at NASA Glenn provided the needed benchmarking data for this project. Special thanks also go to Fred Oswald for his assistance to measure dynamic gear tooth forces. I thank the staff of the Timken Company for their hospitality and for the discussions that took place at their facility during the beginning phase of the project. I especially thank J. David Cogdell for providing surface inspections to support this project, including the surface maps that appear in chapter 4. I thank the staff of the Gear Research Center, University of Toledo, for providing access to tools that furthered the research presented herein. I thank the center's director, Dr. Ahmet Kahrahman, for useful discussions and coordination of activities. I thank Carmen Cioc for providing the executable code, examples, and needed discussions that enabled the lubrication analyses contained herein. I also thank Laurentiu Moraru for the time, effort, and expertise to complete the surface inspections that are included in chapter 7. I received excellent training in fundamentals of statistics via a set of courses taught by Dennis Keller. Mr. Keller was also gracious with his time beyond the classroom setting. Unfortunately, the acknowledgement of his contributions comes posthumous. Mr. Keller's ideas and advice helped greatly to set forth the proper approach for this project.

Available from

NASA Center for Aerospace Information
7121 Standard Drive
Hanover, MD 21076

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22100

Available electronically at <http://gltrs.grc.nasa.gov>

Contents

Chapter 1—Introduction	
1.1 Background.....	1
1.2 Scope and Organization of Document.....	3
Chapter 2—Fundamentals of Statistical Methods for Fatigue Data	
2.1 Introduction	5
2.2 Concepts for Statistical Inference.....	5
2.3 The Weibull Distribution.....	6
2.4 Confidence Intervals for Sample Statistics From Weibull Distributions.....	10
2.5 Discussion.....	11
Chapter 3—Statistics for Gear Fatigue Data (Evaluation, Implementation, and Developments)	
3.1 Introduction	17
3.2 Methods to Estimate Parameters of a Weibull Distribution-Implementation.....	18
3.3 Simulating Effects of Random Sampling on Weibull Sample Statistics	19
3.4 Three Methods for Estimating Weibull Sample Statistics-Qualitative Assessment.....	22
3.5 A Methodology for Quantitative Assessment of Statistical Procedures.....	23
3.6 Presumption of a Zero-valued Threshold Parameter—A Critical Review	25
3.7 Influence of the Number of Samples on the Accuracy and Precision of 10-Percent Life Estimates.....	28
3.8 Influence of Censoring on the Accuracy and Precision of 10-Percent Life Estimates.....	29
3.9 A Summary of the Assessments of Statistical Methods and a Recommendation.....	31
3.10 Confidence Intervals—Implementation.....	32
3.11 A New Method for Comparing Two Datasets to Assess Life Improvements	33
3.12 Conclusions and Recommendations.....	35
Chapter 4—Experimental Evaluation of Gear Surface Fatigue Life	
4.1 Introduction	83
4.2 Test Apparatus, Specimens, and Procedure.....	84
4.3 Results and Discussion	86
4.4 Conclusions	88
Chapter 5—Evaluation of the Experimental Conditions (Methodology)	
5.1 Introduction	101
5.2 Measurement of Dynamic Tooth Loads	101
5.3 Modeling of Residual Stress Profiles	102
5.4 Modeling of Yield Strength Profiles.....	103
5.5 Analysis Method for Determining Contact Pressures of the Lubricated Contact.....	103
5.6 Analysis Method for Determining Sub-Surface Stresses	104
5.7 Analysis Method for Calculation of Load Intensity and Fatigue Damage Indices.....	105
5.8 Summary of Methodology.....	106
Chapter 6—Evaluation of the Experimental Conditions (Results)	
6.1 Introduction	115
6.2 Specific Film Thickness Relation.....	115
6.3 Elastohydrodynamic Analysis	116
6.4 Subsurface Stress Analysis Method and Validation	117
6.5 Stress Analysis Results—Smooth Surfaces.....	118
6.6 Stress Analysis Results—Rough Surfaces.....	122
6.7 Conclusions and Recommendations.....	124
Chapter 7—Gear Tooth Surface Topography	
7.1 Introduction	151
7.2 Description of Inspection Method and Data Display.....	151
7.3 Inspection Results for the As-Manufactured Condition—Ground Gears.....	152

7.4	Inspection Results for the Run-In Condition—Ground Gears.....	152
7.5	Inspection Results for the Run-In Condition—Superfinished Gears.....	154
7.6	Surface Inspections—Summary	155
Chapter 8	—Conclusions	171
Bibliography	174

The Influence of Roughness on Gear Surface Fatigue

Timothy L. Krantz
U.S. Army Research Laboratory
Glenn Research Center
Cleveland, Ohio 44135

Abstract

Gear working surfaces are subjected to repeated rolling and sliding contacts, and often designs require loads sufficient to cause eventual fatigue of the surface. This research provides experimental data and analytical tools to further the understanding of the causal relationship of gear surface roughness to surface fatigue. The research included evaluations and developments of statistical tools for gear fatigue data, experimental evaluation of the surface fatigue lives of superfinished gears with a near-mirror quality, and evaluations of the experiments by analytical methods and surface inspections. Alternative statistical methods were evaluated using Monte Carlo studies leading to a final recommendation to describe gear fatigue data using a Weibull distribution, maximum likelihood estimates of shape and scale parameters, and a presumed zero-valued location parameter. A new method was developed for comparing two datasets by extending the current methods of likelihood-ratio based statistics. The surface fatigue lives of superfinished gears were evaluated by carefully controlled experiments, and it is shown conclusively that superfinishing of gears can provide for significantly greater lives relative to ground gears. The measured life improvement was approximately a factor of five. To assist with application of this finding to products, the experimental condition was evaluated. The fatigue life results were expressed in terms of specific film thickness and shown to be consistent with bearing data. Elastohydrodynamic and stress analyses were completed to relate the stress condition to fatigue. Smooth-surface models do not adequately explain the improved fatigue lives. Based on analyses using a rough surface model, it is concluded that the improved fatigue lives of superfinished gears is due to a reduced rate of near-surface micropitting fatigue processes, not due to any reduced rate of spalling (sub-surface) fatigue processes. To complete the evaluations, surface inspections were completed. The surface topographies of the ground gears changed substantially due to running, but the topographies of the superfinished gears were essentially unchanged with running.

Chapter 1—Introduction

1.1 Background

The subject of this research project is gear surface fatigue, with special attention given to the influence of surface roughness. Gear teeth working surfaces are subjected to repeated rolling and sliding contacts. For operating conditions common for power transmission applications, the loads are sufficient to cause eventual fatigue of the surface. The surface fatigue capability of a gear is one of the most influential factors that defines the size and weight of a gear, and so this subject is of particular interest and importance to the field of aircraft design. This research project sought to provide experimental data and analytical tools to further the understanding of the causal relationship of gear surface roughness to surface fatigue.

The subject of this research came about from the author's involvement with an experimental evaluation of gear surface fatigue. As the time came to select a dissertation topic with specific objectives, preliminary experimental results were showing great potential for improving gear fatigue lives by way of improving surface finish. However, the reason why the improved surface finish provided for an improved fatigue life was not fully understood by the technical community, and it was realized that tools needed to apply the laboratory evaluations to practical engineering applications were lacking. Thereby, the topic for dissertation research was selected to be gear surface fatigue with special attention to the influence of surface roughness. Three broad objectives were set forth: (1) to conduct gear surface fatigue experiments in a controlled manner to provide a quantitative assessment of the relation of surface finish to fatigue life; (2) to provide statistical tools needed to describe and make statistical inference about the fatigue test results; and (3) to conduct analytical investigations to provide a qualitative understanding for the improved fatigue performance. More specific objectives are listed in the individual chapters of this document.

1.2 Scope and Organization of Document

This document consists of this introductory chapter, a final summary chapter, and six main chapters (chapters 2 to 7). Each chapter includes an introductory section, a review of appropriate literature, and a summary of findings, recommendations, or conclusions. Each of the main chapters will now be described in turn.

Chapter 2 provides for a review of concepts and methods for statistics as apply to the present work. The concepts and existing methods are described in some detail providing a concise tutorial of statistics for fatigue data. The chapter notes several instances of conflicting advice and unresolved issues, and specific issues to be resolved by the present research are defined.

The subjects of chapter 3 are the assessments and developments of statistical methods for fatigue data. The work done resolved the specific issues listed in chapter 2. Assessments of statistical procedures were done making use of Monte Carlo studies. The chapter includes a validation study of the Monte Carlo method. Methods for fitting fatigue data to the statistical distribution of choice (the Weibull distribution) were compared and evaluated for accuracy and precision. The evaluations were done giving considerations to: (1) the number of samples usually available, and (2) the use of censoring (suspensions) for fatigue testing. The usual approach of describing the data using a 2-parameter Weibull distribution rather than the more general 3-parameter Weibull distribution was critically examined. The end result of the examination is a recommendation to make use of the 2-parameter Weibull distribution and to fit the parameters using the maximum likelihood method. Next, methods for calculation of confidence intervals are assessed by review of the literature, and a likelihood-based method is selected as the method of choice for the present work. Lastly, a new method is proposed for comparing two gear fatigue datasets. The proposed method is an extension of likelihood-based statistics. The new method is defined and illustrated by an example.

The subject of chapter 4 is an experimental evaluation of the causal relation of surface finish to gear fatigue life. The experiments offer evidence that that gears with differing as-manufactured surface topographies can have dramatically differing performance characteristics. Gear test specimens were prepared having a mirror-like quality surface, a better quality than the usual ground-gear surface finish for aircraft and other vehicles. The method used to provide the mirror-like surfaces is known in the industry as “superfinishing,” and in this document the gears with a mirror-like tooth surface finish will be called “superfinished gears”. The gear specimens, lubrication conditions, load, and speeds were selected such that the test results of the present work could be compared to the NASA Glenn gear fatigue database. The experiments provide for both a qualitative and quantitative measure of the improved fatigue performance that can be provided by superfinishing gears. The statistical methods developed in chapter 3 are applied to describe the test results and to quantify performance differences relative to ground gears. The text of chapter 4 in this document is a minor revision of the peer-reviewed article “Surface Fatigue Lives of Case-Carburized Gears With an Improved Surface Finish,” *Transactions of the ASME, Journal of Tribology*, vol. 123, no. 4.

To apply the performance improvements that were demonstrated by laboratory evaluations to products in the field requires engineering understanding, analysis and judgment. In chapter 5, methodologies are developed for evaluating the experimental condition and, thereby, help provided the tools and data needed for applying the laboratory evaluations. Special experiments were conducted to measure the dynamic forces on the gear teeth during fatigue testing. Next, existing experimental data were used to model the residual stresses and yield strengths as a function of depth below the case-carburized tooth surface. The dynamic loads, residual stresses, and yield strength data were included as part of a contact analysis and assessment of the stress condition of the concentrated line contacts. The contact analysis was done giving consideration to the lubrication condition. Lubrication modeling was done using a computer code developed by others and made available for the present project. A general numerical method was selected, and a computer code developed, to calculate the stress condition for any arbitrary contact pressure distribution. In this manner, rough surface contacts were analyzed. Lastly, a methodology was developed to assess the load intensity as relates to contact fatigue using three alternative stress-based indices.

The subject of chapter 6 is the evaluation of the experimental condition of chapter 4, making use of the methodologies of chapter 5. A series of evaluations are made using progressively fewer assumptions. First, the test results of the current work are presented in terms of a lubrication film thickness-to-roughness ratio, and the data are compared to results of another researcher. Although the film thickness-to-roughness ratio has proven to be a useful concept, it is shown that for the present work further evaluations are warranted. Next, the test results are evaluated, using the methods of chapter 5, assuming that the surfaces are ideally smooth. The stress condition is evaluated in detail. The possibility that fatigue life improvements are the result of reductions in friction due to the superfinishing is investigated. Reductions in friction do not seem to be the primary effect for the fatigue life differences. Lastly, the experimental conditions are evaluated modeling the surfaces as rough surfaces. The evaluations using rough surface models provide qualitative assessments of the experimental condition.

The title of chapter 7 is “Gear Tooth Surface Topography.” Toward the latter part of this research project, an interferometric microscope inspection machine became available to this investigation. Some of the tooth surfaces for this project were inspected, initially out of curiosity. The inspections revealed many interesting features. It was decided that even though the inspections did not resolve any issues, the inspection data was a valuable contribution to the literature. Additional inspections were made, and a number of these inspections are organized, presented, and discussed in chapter 7.

Chapter 8 is a final summary of the project. Final conclusions and specific contributions made to the state-of-the-art are summarized.

Chapter 2—Fundamentals of Statistical Methods for Fatigue Data

2.1 Introduction

For certain applications, gears are designed to operate for many cycles with a high degree of probability for survival. For example, the design criteria for aircraft may require that 90-percent of the gears will survive at least 10^9 revolutions without fatigue failure. One potential failure mechanism is surface fatigue of the contacting gear tooth surfaces. For the example design criteria stated above, the gears will be subject to the possibility of high cycle fatigue failure. Consistent with other types of high cycle fatigue phenomena, the surface fatigue lives of nominally identical gears operated in a nominally identical fashion will vary greatly from one specimen to the next. Therefore, statistical concepts and methods are important to effectively evaluate and make use of data from surface fatigue experiments. The section to follow that describes statistical concepts draws heavily from the text of Meeker and Escobar (ref. 2.1).

2.2 Concepts for Statistical Inference

Laboratory evaluations of fatigue life can be considered as analytical studies as defined by Deming (ref. 2.2). Analytical studies answer questions about processes that generate output over time. For analytical studies, one makes predictions about future behavior (inference) by analysis of past behavior, and the results strictly apply only to an unchanging process. All processes will change, to some extent, over time. Therefore, the size of a statistical interval obtained from an analytical study must be considered as a lower bound on the precision with which one can predict future behavior. Furthermore, in analytical studies the process from which samples are obtained for evaluation may differ from the target process of interest. For example, production lines are often not available for manufacturing prototypes. In such an example, statistical methods can directly quantify the future behavior only of products made in the same manner as the prototypes, and so one must use engineering judgment and experience to make predictions about the future behavior of products from a production line process.

Although gear surface fatigue data are of a discrete nature (integer number of cycles), the distributions of the data are usually modeled using continuous scales. For an analytical study, which is the interest of this work, the cumulative distribution function, defined by

$$F(n) = \Pr(N \leq n), \quad (2.2.1)$$

gives the probability that a unit will fail within n load cycles. Alternatively, the cumulative distribution function can be interpreted as the proportion of units taken from a stationary (unchanging) process that will fail within n load cycles. The probability density function is the derivative of $F(n)$ with respect to n ,

$$f(n) = \frac{dF(n)}{dn}. \quad (2.2.2)$$

For a normal distribution the probability distribution function has the shape of the familiar bell curve. The hazard function expresses the propensity for a system to fail in the next load cycle. It is related to the cumulative distribution and probability density functions as

$$h(n) = \frac{f(n)}{[1 - F(n)]}. \quad (2.2.3)$$

Fatigue testing often employs the concept of censoring. In this work, it is assumed that all tests have either been run to failure or have been suspended (censored), without failure, at a prespecified time. This

type of censoring scheme is commonly known as Type I censoring. Effective use of censoring can enable one to arrive at statistical conclusions with less total test time than running all units to failure.

The concept of the sampling distribution (ref. 2.3) is a key concept for making statistical inference. The concept will be described by an example. Consider that one is interested in determining the value of the cumulative distribution function of a process for a certain number of cycles, $F(N)$. To estimate the value, a finite number of random samples are obtained, tested, and analyzed to determine an estimate, $\tilde{F}(N)$. The standard approach to quantify the possible size of difference between the true but unknown value, $F(N)$, and the estimate, $\tilde{F}(N)$, is to consider what would happen if the entire inferential procedure (sampling, testing, and analysis) were repeated many times. Each time, a different estimate, $\tilde{F}(N)$, would be obtained since the sets of random samples differ. The distribution of the $\tilde{F}(N)$ values, the sampling distribution, provides insight about the true value, $F(N)$. The spread of the sampling distribution is often called the “sampling error.” The sampling distribution is a function of the cumulative distribution function, the number of samples, and the procedure for estimating the value of the function.

A common way to quantify uncertainty due to “sampling error” is by the use of confidence intervals. Confidence intervals are stated using some specified level of confidence. The level of confidence describes the performance of a confidence interval procedure and, thereby, expresses one’s confidence that a particular interval contains the quantity of interest. A summary of methods for calculation and interpretation of commonly used intervals is available (ref. 2.4). In some cases, confidence intervals can be defined by exact analytical expressions. In other cases, one must employ approximate methods. For approximate methods, the stated confidence only approximates the true confidence that the interval contains the quantity of interest.

Both parametric and nonparametric methods have been developed for making statistical inference. Nonparametric methods do not require that the analyst make any assumptions about the form of the failure distribution, whereas parametric methods require such an assumption. Noting another difference, nonparametric methods require reporting all of the data and/or a graphical representation of the data, while parametric models allow for complete descriptions of datasets by defining a few parameters. Parametric models provide for smooth estimates of failure-time distributions and make possible extrapolations into the tails of the distribution. In general, confidence intervals for parametric models are smaller than the same intervals for nonparametric models. The advantages of parametric models relative to nonparametric models carry with them the assumption that the proposed parametric form is appropriate. Serious errors can arise if the assumption is not valid. In this work, it is assumed that the failure-time distributions for gear surface fatigue are adequately modeled by a distribution known as the Weibull distribution, the next topic for discussion.

2.3 The Weibull Distribution

The Weibull distribution has been presented in early work by the developer, W. Weibull, as one that provided reasonable descriptions for a wide variety of phenomena (ref. 2.5). It has since been used in many fields of study, and it is now widely accepted as a parametric model for reliability and fatigue data. From the theory of extreme values, one can show that the Weibull distribution models the minimum of a large number of independent positive random variables from a certain class of distributions. This relationship to extreme value statistics has provided a framework for studying the properties of the distribution and has provided some theoretical basis for its application to fatigue data. A common justification for its use is empirical: it can model failure-time data regardless of whether the hazard function is increasing (appropriate for cumulative wear phenomena) or decreasing (appropriate for infant mortality phenomena), and the probability density function may be either skewed left or skewed right.

Several formulas, different in detail but mathematically equivalent, have appeared in the literature as the mathematical definition of the Weibull distribution. Unfortunately, the differing formulas sometimes share common terminology, and this situation has been the source of some confusion (ref. 2.6). In this work, the following definition of the cumulative distribution function from reference 2.1 has been adopted,

$$F(t) = 1 - \exp \left[- \left(\frac{t - \gamma}{\eta} \right)^\beta \right], \quad (2.3.1)$$

where γ is the threshold parameter, β is the shape parameter, and η is the scale parameter. The shape of the probability density function is determined by the value of the shape parameter (fig. 2.3.1). Typically, gear and bearing surface fatigue data have distributions with shape parameters between 1 and 3, and therefore the probability distribution functions are typically skewed right. For a given Weibull distribution, one can determine the mode, mean, median, variance, and other properties by employing the gamma function (ref. 2.7)

Methods to Estimate Parameters of Weibull Distributions.—With the adoption of the Weibull distribution to a wide variety of phenomena, the technical community has shown great interest in the development of methods for estimating the distribution parameters from sample data. The study of methods for estimation continues today. For example, reference 2.8 describes a study comparing eight different methods. There exist at least four broad classes of estimators: graphical methods, method of moments, regression methods, and likelihood-based methods. Nelson (ref. 2.9) wrote a classic article defining a graphical method for the analysis of failure data using the concept of the hazard function. Later, Nelson suggested that such graphical methods, which require fitting of data by eye, should be complemented by analytical methods (ref. 2.10). The method of moments is widely used for fitting distributions, but the method is not appropriate for censored data (ref. 2.11) and so will not be considered in the present study. Regression methods and likelihood methods, both considered in the present work, will be described in some detail in the text to follow.

Weibull distribution parameters can be estimated from sample data using linear least-squares regression by making a transformation of the cumulative distribution function (eq. (2.3.1)). Taking the natural logarithm of the natural logarithm of both sides and simplifying provides the needed transformation as

$$\ln \left\{ \ln \left[\frac{1}{1 - F(t)} \right] \right\} = \beta \ln(t - \gamma) - \beta \ln(\eta). \quad (2.3.2)$$

Equation (2.3.2) is in a form appropriate for linear least-squares regression, but closed form expressions to determine the three parameters that minimizes the sum of the squared-errors are not available. Iterative methods (ref. 2.12, for example) must be used to estimate the threshold parameter, γ . For many applications, the threshold parameter is assumed known, and a common value assumed is zero. If the threshold parameter is assumed known, then closed form expressions are used to minimize the sum of the squared-errors of the transformed dependent variable. The case of an assumed zero-valued threshold parameter is sometimes called the 2-parameter Weibull distribution.

For many applications, the linear least-squares regression method is a well-defined one. However, there exist in the literature significant differences in the details of its application to the fitting of data to the Weibull distribution. As an example of one difference, weighting functions have been proposed and used by some researchers (refs. 2.8 and 2.13). As for a second difference, some authors have considered the dependent variable to be the observed (transformed) sample times to failure (ref. 2.14) while others have considered the dependent variable to be the (transformed) cumulative failure probability (ref. 2.6). Regardless of the selection of the dependent variable, one must assign a cumulative failure probability to each failure in the dataset. This leads us to a discussion of a third difference, as there exist in the literature several methods for assigning the cumulative failure probability (or the so-called “plotting position”). For all methods, the observed failure times are first ranked from smallest to largest. The question concerns the assignment of a cumulative probability for ranked observation “ i ” from a total of “ n ” observations. The simple proportion formula known as the Kaplan-Meier relation,

$$p(i, n) = \frac{i}{n} \quad (2.3.3)$$

is sometimes suggested (ref. 2.6) even though it is known to have deficiencies. Several other very simple formulae are available. The mean rank relation,

$$p(i, n) = \frac{i}{n + 1} \quad (2.3.4)$$

is sometimes used (ref. 2.15), but it is known to provide biased estimates of the shape parameter (ref. 2.13). A mid-point plotting position that is suggested by several authors (refs. 2.1, 2.10, and 2.13) is defined by the relation

$$p(i, n) = \frac{i - 0.5}{n}, \quad (2.3.5)$$

and reference 2.16 reports a standard that calls for using a modified version of the mid-point formula

$$p(i, n) = \frac{i - 0.5}{n + 0.25}. \quad (2.3.6)$$

The Nelson estimator, derived from nonparametric concepts, is given as (ref. 2.16)

$$p(i, n) = 1 - \exp \left[- \sum_{j=1}^i \left(\frac{1}{n - j + 1} \right) \right]. \quad (2.3.7)$$

Perhaps the most widely used estimators of cumulative probability are those that are based on the idea of providing median rank positions. The philosophy of the median rank is to provide estimates that will be too large 50 percent of the time and, therefore, are also too small 50 percent of the time. An approximate equation for median rank has been established (ref. 2.17) as

$$p(i, n) = \frac{i - 0.3}{n + 0.4}. \quad (2.3.8)$$

Analytical expressions for the exact values for median ranks are known, but the equations require the numeric evaluation of the incomplete beta function. Tables of median rank values have been compiled (refs. 2.14 and 2.18). Jacquelin (ref. 2.19) provides a robust algorithm for calculating exact median rank values.

Alternatives to regression methods for fitting sample data to parametric distributions have been developed. A very popular method is the maximum likelihood method. The likelihood function can be described as being proportional to the probability of the data. The total likelihood of a set of data equals the joint probability of the individual data points. Assuming n independent observations, the sample likelihood is (ref. 2.1)

$$L(\beta, \gamma, \eta; (x_1, x_2, \dots, x_n)) = C \prod_{i=1}^n L_i(\beta, \gamma, \eta; x_i) \quad (2.3.9)$$

where L_i is the probability for observation x_i to occur from a Weibull distribution with parameters β , γ , and η . To estimate parameters β , γ , and η one finds those values that maximizes the total sample likelihood just defined. In usual situations, the constant C in equation (2.3.8) does not depend on the distributional parameters, and so C can simply be taken as $C = 1$ for purposes of parameter estimation. The concept of likelihood is illustrated in figure 2.3.2 showing a Weibull probability distribution function.

Referring to the figure, an interval censored data point is one for which the item was known to have survived at time t_2 , and then failure occurred between times t_3 and t_2 . Using the relationship of the cumulative probability and probability density, the sample likelihood for interval censoring is

$$L_i = \int_{t_2}^{t_3} f(t) dt = F(t_3) - F(t_2). \quad (2.3.10)$$

Left and right censoring can be considered as special cases of interval censoring. Left censoring can occur if failures are observed only at scheduled inspection times and a failure is observed upon first inspection. For left censoring, the sample likelihood is

$$L_i = \int_0^{t_1} f(t) dt = F(t_1) - F(0) = F(t_1). \quad (2.3.11)$$

Right censoring occurs if an item survives without failure. For example, fatigue testing might be suspended at a prespecified time, and if the test has not produced fatigue at such a time the result is a right-censored data point. Special care must be exercised in the treatment of field data. For example, specimens removed without failure, but for cause, (because of some sign of impending fatigue) should not be considered as right-censored data. For right censoring, the sample likelihood is

$$L_i = \int_{t_4}^{\infty} f(t) dt = F(\infty) - F(t_4) = 1 - F(t_4). \quad (2.3.12)$$

Strictly speaking, data reported as exact failure times are truly interval censored since the data are of finite precision. However, an approximation often used for the sample likelihood is

$$L_i = f(t_i) \propto f(t_i) \Delta_i \approx [F(t_i) - F(t_i - \Delta_i)], \quad (2.3.13)$$

where Δ_i is small and does not depend on the distributional parameters. Since the density approximation is proportional to the true likelihood for appropriately small Δ_i then the shape of the likelihood function and the location of the maximum is not affected. This approximation is often employed since it can yield, in some cases, closed form expressions for maximum likelihood estimates of the distributional parameters. The approximation is not always adequate. If the density approximation is used for fitting a 3-parameter Weibull distribution, then for some datasets the total likelihood can increase without bound in the parameter space. Hirose (ref. 2.20) suggests that one can use a generalized extreme value distribution as an extension of the Weibull distribution to assess such a troublesome situation.

Often, equations involving likelihood can be simplified if the log-likelihood is employed. The total log-likelihood is defined in terms of a sum as

$$LL = \log[L] = \sum_{i=1}^n \log(L_i) \quad (2.3.14)$$

If a maximum exists for the log-likelihood, then the maximum will occur for the same parameter values as will the maximum for the likelihood.

The maximum likelihood estimates for the parameters of a distribution are those that maximize the total likelihood of the data. For an assumed Weibull distribution form, analytical expressions have been developed in the usual way, by taking derivatives with respect to the unknown parameters and then setting the expressions for the derivatives to zero (refs. 2.21 and 2.22). The parameter values that

simultaneously satisfy all expressions must be found iteratively. Keats, Lawrence, and Wang (ref. 2.23) published a Fortran routine for maximum likelihood parameter estimation of two-parameter Weibull distributions.

Even though maximum likelihood estimates (MLE) of Weibull parameters were introduced more than 35 years ago, the properties of MLE continue to be studied and characterized today. It has been established that MLE of Weibull distributions can be biased. Here, the word biased indicates that the true value of the parameter and the mean of the sampling distribution of the MLE of that parameter are not equal. Various bias correction methods have been proposed and studied (refs. 2.24, 2.25, and 2.26). Jacquelin (ref. 2.27) has provided tables of correction factors. McCool (ref. 2.28) proposed and developed methods for point estimates with median unbiased properties. Cacciari, et al. (ref. 2.24) contend that the bias is not the only important characteristic, and they caution that in certain situations the unbiasing methods are inappropriate.

2.4 Confidence Intervals for Sample Statistics From Weibull Distributions

The most common way for one to quantify uncertainty due to “sampling error” is to state a confidence interval. Confidence intervals are stated using some specified level of confidence. The level of confidence describes the performance of a confidence interval procedure and, thereby, expresses one’s confidence that a particular interval contains the quantity of interest. A summary of methods for calculation and interpretation of commonly used intervals is available (ref. 2.4). For samples from normal distributions with no censoring, the technical community has established mathematically rigorous, exact confidence interval methods. However, for Weibull distributions special methods are often required to calculate confidence intervals. For Type I censoring, exact methods for intervals are not available (ref. 2.29). Because gear fatigue testing typically employs Type I censoring, only approximate methods were investigated and considered for this project. Confidence interval methods can be classified into three groups (ref. 2.30). One group considers that the sampling distribution of the sample statistic is, approximately, a normal distribution. The sample statistic may be a transformed value (for example, the log of the sample shape parameter). A second group of confidence interval methods is based on likelihood ratio statistics and modifications of those statistics. The third group of confidence interval procedures is based on parametric bootstrap methods (ref. 2.31) that make use of Monte Carlo simulations. Some basic concepts for these three groups will be described in turn.

Confidence intervals based on normal-approximation theory are perhaps the most widely used intervals. Most commercial software packages calculate intervals by this method (ref. 2.30). The method is based on asymptotic theory and holds well for large sample sizes. For maximum likelihood estimators, the normal distribution approximation can be improved by first providing an appropriate transformation. Meeker and Escobar (ref. 2.1) provide detailed equations and examples. Studies have shown (refs. 2.29 and 2.30) that even with appropriate transformations, the asymptotically normal methods converge to nominal error probabilities rather slowly and, therefore, perform poorly for small sample sizes.

Likelihood ratio statistics are used as another method for confidence intervals. Using a two-parameter Weibull distribution as an example, the profile likelihood function for the shape parameter β is

$$R(\beta) = \max_{\eta} \left[\frac{L(\beta, \eta)}{L(\hat{\beta}, \hat{\eta})} \right]. \quad (2.4.1)$$

In the previous equation, for a fixed value of β , η is determined to provide the maximum value for the ratio. The denominator of the ratio is the likelihood value found using the maximum likelihood estimates of the parameters. Once the profile likelihood is determined, an approximate confidence interval for 100(1- α) percent confidence is

$$R(\beta) > \exp\left[-\chi^2_{(1-\alpha;1)} / 2\right] \quad (2.4.2)$$

As the method is based on likelihood, it cannot be applied to estimators found by regression.

A more recently developed technique for calculating confidence intervals is the parametric bootstrap method introduced by Efron (ref. 2.31). Monte Carlo simulation is used to determine the properties of a particular confidence interval. This method can be applied so long as the inference procedure is well defined and automated with a robust algorithm. It is particularly fitting for complicated censoring schemes. McCool (ref. 2.28) developed test statistics for Weibull populations that depend on the sample size and number of failures but not on the values of the Weibull parameters, and he established confidence intervals for such statistics using Monte Carlo simulation.

As a final comment on intervals, work has been done to develop prediction interval methods for Weibull distributions from censored sample data (refs. 2.32 and 2.33). Hahn and Meeker (ref. 2.32) describe the difference between confidence and prediction intervals. The present work makes use of confidence intervals.

2.5 Discussion

A practitioner faced with the task of statistical analysis of fatigue and reliability data needs to select the most appropriate method from the many that have been proposed. The preceding sections of this chapter highlight some of the open questions concerning estimating parameters of a Weibull distribution from sample data. The interest of this work is narrower than many of the references cited in this work. For gear surface fatigue data, probability density functions are typically skewed right, samples sizes usually range from 10 to 40, and often censoring is limited to Type I censoring. In addition, the estimation of the distribution parameters is often a means to an end, the goal being the estimation of percentiles of the cumulative distribution function.

Studies have been completed to provide guidance for the statistical treatment of gear surface fatigue data. The results of the appropriately focused studies will be described in chapter 3. Four specific issues were resolved.

1. Software for the calculation of parameter estimates and confidence intervals were developed and validated.
2. Three methods for determining distribution parameters from sample data (two regression-based methods and a maximum likelihood-based method) were evaluated for accuracy and precision.
3. The usual practice of describing gear surface fatigue data using the 2-parameter Weibull distribution rather than the more general 3-parameter Weibull distribution was critically examined.
4. A new method is proposed for comparing two datasets. The new method was developed to detect the existence of statistically significant differences in fatigue life properties. The method compares two datasets based on a selected quantile of the fatigue life cumulative distribution functions.

References

- 2.1. Meeker, W.; Escobar, L., *Statistical Methods for Reliability Data*, John Wiley and Sons, New York, 1998.
- 2.2. Deming, W., "On Probability as a Basis For Action," *The American Statistician*, 1975.
- 2.3. Fisher, R., "The General Sampling Distribution of the Multiple Correlation Coefficient," *Proceedings of the Royal Society of London, Ser. A*, **121**, (1928).
- 2.4. Hahn, G.; Meeker, W., *Statistical Intervals – A Guide for Practitioners*, John Wiley and Sons, New York, 1991.
- 2.5. Weibull, W., "A Statistical Distribution of Wide Applicability," *J. Appl. Mech.*, **18**, (1951).
- 2.6. Hallinan, A., "A Review of the Weibull Distribution," *J. of Quality Technology*, **25** [2], (1993).

- 2.7. Cohen, A., "The Reflected Weibull Distribution," *Technometrics*, **15** [4], (1973).
- 2.8. Montanari, G. et al., "In Search of Convenient Techniques for Reducing Bias in the Estimation of Weibull Parameters for Uncensored Tests," *IEEE Trans. On Dielectrics and Electrical Insulation*, **4** [3], (1997).
- 2.9. Nelson, W., "Theory and Applications of Hazard Plotting for Censored Failure Data," *Technometrics*, **14**, (1972).
- 2.10. Nelson, W., *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, John Wiley and Sons, New York (1990).
- 2.11. *Electronic Statistics Textbook*, Tulsa, Oklahoma: StatSoft, Inc., (1999). (Available online: <http://www.statsoft.com/textbook/stathome.html>) Accessed August 6, 2001.
- 2.12. Soman, K.; Misra, K., "A Least Square Estimation of Three Parameters of a Weibull Distribution," *Microelectronics Reliability*, **32** [3], (1992).
- 2.13. Bergman, B., "Estimation of Weibull Parameters Using a Weight Function," *Journal of Materials Science Letters*, **5** [6], (1986).
- 2.14. Abernathy, R.; Breneman, J.; Medlin, C.; Reinman, G., *Weibull Analysis Handbook*, Air Force Wright Aeronautical Laboratories Report AFWAL-TR-83-2079, (1983). Available from National Technical Information Service, Washington, DC.
- 2.15. Pieracci, A., "Parameter Estimation for Weibull Probability Distribution Function of Initial Fatigue Quality," *AIAA Journal*, **33** [9], (1995).
- 2.16. Cacciari, M.; Montanari, G., "Discussion - Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and other Probability Paper," *IEEE Trans. On Electrical Insulation*, **26** [6], (1991).
- 2.17. Fothergill, J., "Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and other Probability Paper," *IEEE Trans. On Electrical Insulation*, **25** [3], (1990).
- 2.18. Johnson, L., *The Statistical Treatment of Fatigue Experiments*, Elsevier, New York, (1964).
- 2.19. Jacquelin, J., "A Reliable Algorithm for the Exact Median Rank Function," *IEEE Trans. On Electrical Insulation*, **28** [2], (1993).
- 2.20. Hirose, H., "Maximum Likelihood Estimation in the 3-parameter Weibull Distribution," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996) and ERRATUM, **3** [2], (1996).
- 2.21. Harter, H.; Moore, A., "Maximum-Likelihood Estimation of the Parameters of Gamma and Weibull Populations from Complete and from Censored Samples," *Technometrics*, **7** [4], (1965).
- 2.22. Cohen, A., "Maximum Likelihood Estimation in the Weibull Distribution Based On Complete and On Censored Samples," *Technometrics*, **7** [4], (1965).
- 2.23. Keats, J.; Lawrence, F.; Wang, F., "Weibull Maximum Likelihood Parameter Estimates with Censored Data," *J. of Quality Tech.*, **29** [1], (1997).
- 2.24. Cacciari, M.; Mazzanti, G.; Montanari, G., "Comparison of Maximum-Likelihood Unbiasing Methods for the Estimation of the Weibull Parameters," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996).
- 2.25. Hirose, H., "Bias Correction for the Maximum Likelihood Estimates in the Two-parameter Weibull Distribution," *IEEE Trans. On Dielectrics and Electrical Insulation*, **6** [1], (1999).
- 2.26. Ross, R., "Bias and Standard Deviation due to Weibull Parameter Estimation for Small Data Sets," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996).
- 2.27. Jacquelin, J., "Inference of Sampling on Weibull Parameter Estimation," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [6], (1996).
- 2.28. McCool, J., "Evaluating Weibull Endurance Data by the Method of Maximum Likelihood," *ASLE Transactions*, **13** [3], (1970).
- 2.29. Doganaksoy, N.; Schmee, J., "Comparisons of Approximate Confidence Intervals for Distributions Used in Life-Data Analysis," *Technometrics*, **35** [2], (1993).
- 2.30. Jeng, S.; Meeker, W., "Comparisons of Approximate Confidence Interval Procedures for Type I Censored Data," *Technometrics*, **42** [2], (2000).
- 2.31. Efron, B.; Tibshirani, R., *An Introduction to the Bootstrap*, Chapman & Hall, New York (1993).

- 2.32. Engelhardt, M.; Bain, L., “On Prediction Limits for Samples From a Weibull or Extreme-value Distribution,” *Technometrics*, **24** [2], (1982).
- 2.33. Escobar, L., Meeker, W., “Statistical Prediction Based on Censored Life Data,” *Technometrics*, **41** [2], (1999).

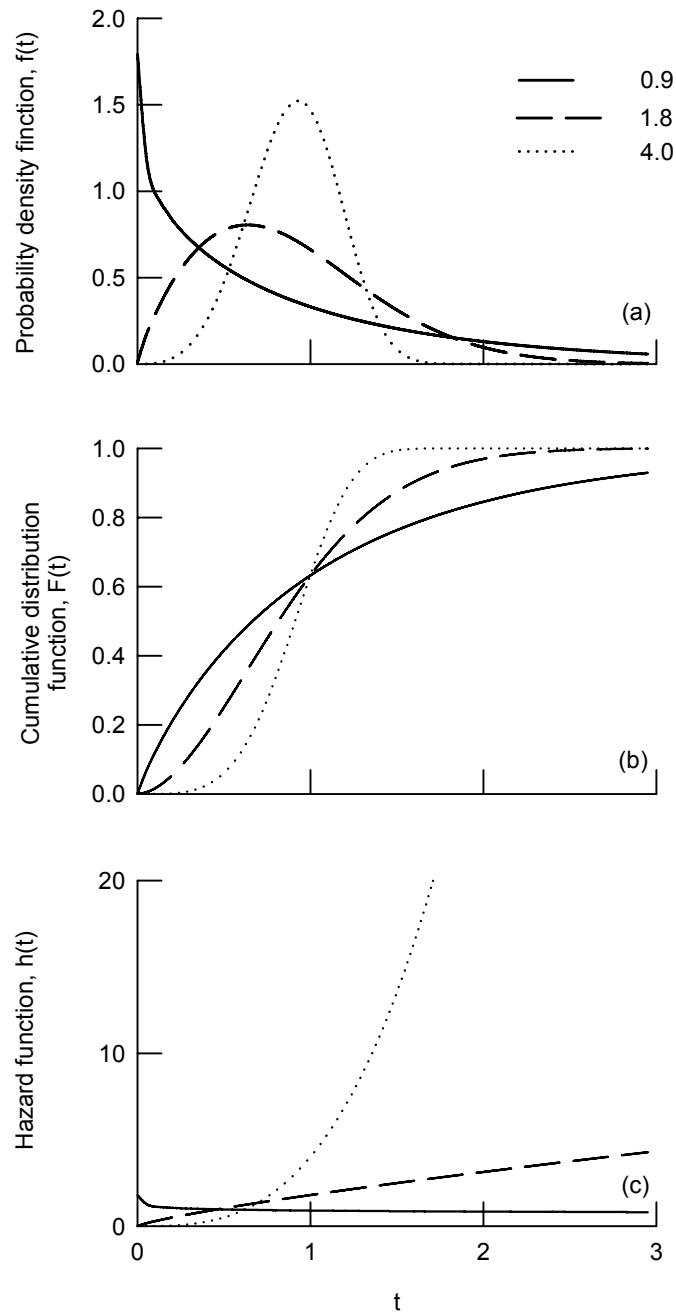


Figure 2.3.1.—Weibull distributions for three shape factors.
 (a) Probability density functions. (b) Cumulative density functions. (c) Hazard functions.

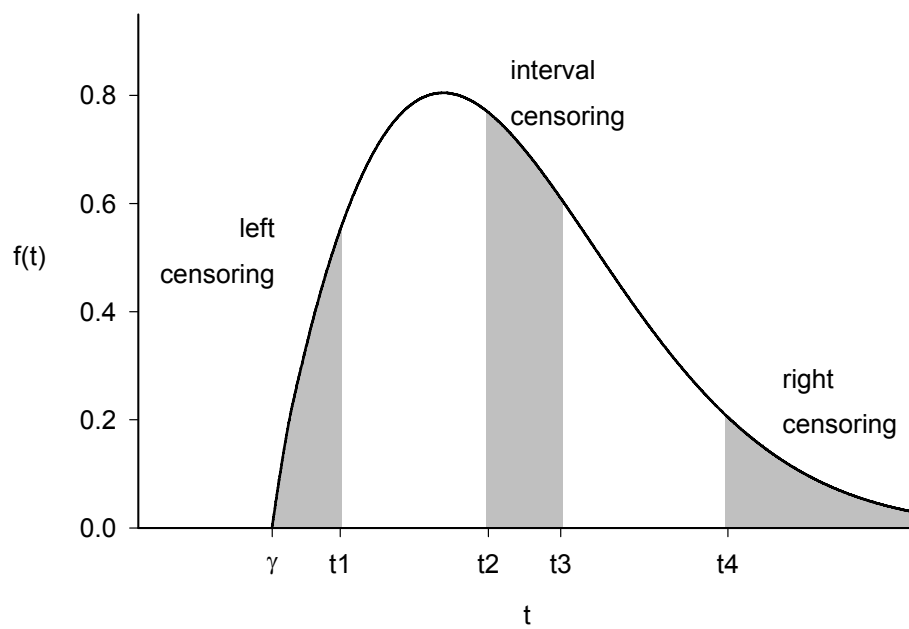


Figure 2.3.2.—Likelihood contributions for three types of censoring. In the manner of Meeker and Escobar (ref. 2.1).

Chapter 3—Statistics for Gear Fatigue Data (Evaluation, Implementation, and Developments)

3.1 Introduction

Consistent with other types of high cycle fatigue phenomena, the surface fatigue lives of nominally identical gears operated in a nominally identical fashion will vary greatly from one specimen to the next. Therefore, statistical concepts and methods are important to effectively evaluate and make use of data from gear surface fatigue experiments. Some of the basic concepts and methods for statistical analysis of fatigue data were presented in chapter 2. The final section of chapter 2 discusses some of the conflicting advice and open issues concerning the application of statistical analysis to gear surface fatigue data.

A practitioner faced with the task of statistical analysis of fatigue and reliability data needs to select the most appropriate method from the many that have been proposed. For gear surface fatigue data, probability density functions are typically skewed right, samples sizes usually range from 10 to 40, and often censoring is limited to Type I censoring. In addition, the estimation of the distribution parameters is often a means to an end, the goal being the estimation of quantiles of the cumulative distribution function. Another end goal is to be able to compare two populations with (potentially) differing fatigue life distributions. With these ideas and end goals in mind, appropriately focused studies and developments were completed. Those studies, and the recommendations coming from those studies, are the subject of this chapter.

There exists conflicting advice concerning the preferred method for the statistical description and inference of fatigue data. In this work, three methods for describing gear fatigue life distributions are implemented and assessed to determine which of the three is preferred for typical gear surface fatigue data. The first step of the study was to implement and then validate the algorithms for all three estimating methods (section 3.2). The primary tool used to evaluate and compare the three estimating methods was the Monte Carlo simulation of random sampling effects. Details of the Monte Carlo simulation scheme and validation studies are provided in section 3.3. The Monte Carlo tool is then used to assess the three estimating methods (sections 3.4 to 3.8). The assessments were done by evaluating the accuracy and precision of sample statistics. The usual presumption of a zero-valued threshold parameter was critically examined and evaluated. The influence of sample size and censoring methods on the precision and accuracy of 10-percent life estimates was also studied. In section 3.9, all of the evaluations and assessments of the three estimating methods are summarized and discussed, and a final recommendation is provided.

Along with a recommendation for a preferred estimating method, tools for statistical inference are implemented and developed. In section 3.10, methods for calculating confidence intervals for sample statistics are reviewed, and a likelihood ratio based method is selected, implemented, and validated. In section 3.11, a new method for comparing two datasets is proposed. The method is a way to compare two datasets that represent two populations with (potentially) differing life distributions. The method is based on the likelihood ratio. The method allows for the comparison based on any chosen quantile. For the present work, the 10-percent life quantile is chosen as the basis for comparison. To employ this new method, a null hypothesis is set forth stating that the 10-percent lives of the two populations are identical. Then, the confidence with which one can reject the null hypothesis (based on the experimental evidence) is calculated. The newly proposed method is illustrated by an example.

The results, conclusions, and recommendations from all of these studies and developments are collected and reported in section 3.12. This final section of the chapter also discusses some ideas, beyond the present scope, for extending the works presented here.

3.2 Methods to Estimate Parameters of a Weibull Distribution—Implementation

In this work, it is assumed that the Weibull distribution is an appropriate one for description of gear fatigue data. Three methods for estimating the parameters of the Weibull distribution from randomly selected samples were selected for study, namely:

1. the 2-parameter least-squares regression method,
2. the 3-parameter least-squares regression method,
3. the 2-parameter maximum likelihood method.

These three methods were selected for study based on review of previous works (chapter 2). This section describes the implementation of these three methods.

The regression-based methods for estimating the Weibull distribution is accomplished by a transformation of the cumulative distribution function to the form of a line.

$$F(t) = 1 - \exp \left[- \left(\frac{t - \gamma}{\eta} \right)^\beta \right], \quad (3.2.1)$$

Although the least-squares regression method is widely used, there exist some differences in the literature concerning details of implementation for the Weibull distribution. Some authors have considered the dependent variable to be the observed (transformed) sample times to failure (ref. 3.1) while others have considered the dependent variable to be the (transformed) cumulative failure probability (ref. 3.2). In this work, the former was selected for implementation. The reasoning is that for a given data point, the time to failure is known (was measured), and so the time-to-failure was considered as the independent variable while the cumulative failure probability was considered as the dependent variable.

Implementation of the regression-based methods requires selecting a method for assigning cumulative failure probabilities. (The cumulative failure probability is also sometimes called the plotting position). Many methods have been proposed and studied in the literature (for examples, see section 2.3). For purposes of this work, the selection of a method for assigning the probabilities was not considered critical. As stated by Nelson (ref. 3.3), “Some authors strongly argue for a particular plotting position. This is as fruitless as arguing religions; they all get you to heaven”. In this work, cumulative failure probabilities are assigned using the exact median ranks. The approach considers that the assigned failure probability will be too high 50-percent of the time and too low 50-percent of the time. Exact median ranks were calculated using a Fortran implementation of the algorithm of Jacquelin (ref. 3.4). Calculations were done to evaluate the accuracy of the following commonly used equation that provides an approximate value for the median rank,

$$p(i, n) = \frac{i - 0.3}{n + 0.4}. \quad (3.2.2)$$

Results of the calculations are provided in figure 3.2.1. The approximate equation provides values that are adequate for most engineering applications. However, since an exact median rank method was readily available, it was implemented and used for the present work.

For the 2-parameter Weibull distribution, the solution of the least-squares estimates of the parameters can be done in the usual manner using closed form expressions. However, the solution for the 3-parameter distribution requires an iterative approach. Recall that the following expression has been adopted as the definition of the 3-parameter Weibull distribution,

$$F(t) = 1 - \exp \left[- \left(\frac{t - \gamma}{\eta} \right)^\beta \right], \quad (3.2.3)$$

where γ is the threshold parameter, β is the shape parameter, and η is the scale parameter. Some authors have studied methods for estimating all three parameters (refs. 3.5 to 3.7). In these cited references, the threshold parameter is constrained to be equal to or greater than zero. On one hand this seems to be a reasonable constraint for fatigue data since the true value for the threshold must be greater than or equal to zero. However, in this work the approach taken is that the *threshold estimate* may take on negative values. If the true threshold value is indeed zero, then for some sets of random samples, the ‘best fit’ will require a negative value for the threshold estimate. Constraining the threshold estimate would introduce an undesired bias into the sampling distribution for the threshold parameter.

To develop an approach for the iterative solution of the least-squares regression for the 3-parameter distribution, Monte Carlo simulation was employed. (See section 3.3 for details of the Monte Carlo implementation). A 3-parameter Weibull distribution was chosen for study having a scale parameter equal to 1.0, a shape parameter equal to 3.0, and a threshold parameter equal to 0.4. Monte Carlo simulation was used to provide 20 psuedo-random samples from this population. From the sample data, plots were created to display the sum of squared errors as a function of presumed values for the threshold parameter. With a threshold value presumed, the shape and scale parameters that minimize the sum of squared errors were found by the usual closed form expressions for regression of a line. Twelve such simulations were completed, and the results are provided in figure 3.2.2. We note that in all cases a single minimum value exists within the range displayed. Therefore, a simple iterative search can be accomplished, within some initially provided bounds, for the threshold parameter. It is noted from figure 3.2.2 that for some datasets, the minimized sum of squared-errors is insensitive over quite a large range of assumed values for the threshold parameter. This previous statement is especially true for the case where the threshold value that provides the optimal solution is negative. Therefore, it is anticipated that the sampling distribution for the threshold parameter will have a large breadth when using the least-squares regression method to estimate the parameters.

Along with the regression-based methods just described, the maximum likelihood method was selected for study. Some background on the method is provided in chapter 2. The maximum likelihood method for the 2-parameter Weibull distribution requires an iterative solution since closed form expression for the parameter values to maximize the likelihood are not available. Keats, Lawrence, and Wang have published a Fortran program to determine the maximum likelihood solution for fitting a 2-parameter Weibull distribution from right-censored data (ref. 3.8). The published program formed the basis of a subroutine used in the present work. The published program was modified to be a subroutine, and all calculations within the subroutine were changed to be double precision calculations, thereby improving the robustness of the solution. The software implementation was validated against solutions published in the open literature before using the routines for further studies (ref. 3.9).

3.3 Simulating Effects of Random Sampling on Weibull Sample Statistics

Background.—The Monte Carlo method is a well-established method for studying the properties of statistical calculation procedures and methods. The method has been used extensively to study the properties of sample statistics for the Weibull distribution (refs. 3.7, 3.10 to 3.18). For sample statistics of interest from the Weibull distribution, the Monte Carlo method has been especially useful since exact analytical expressions for the sampling distributions are unknown. Conceptually, the Monte-Carlo method makes use of a random process to simulate random sampling from a population. Since computerized random number generators use a deterministic algorithm, such generators are more correctly referred to as psuedo-random number generators. Coddington (ref. 3.19) provides suggestions for the use of psuedo-

random number generators, and Jacquelin (ref. 3.14) emphasizes the need for researchers to provide some details about the particulars of one's Monte-Carlo method.

Following the advice of Coddington (ref. 3.19), two random number generators that have been extensively tested were selected for the present work. The first generator makes use of the multiplicative congruential method (refs. 3.20 and 3.21). The second generator makes use of the generalized feedback shift register method (ref. 3.22). Random sampling from a population representing time-to-failure was simulated by generating a set of psuedo-random numbers. The set of numbers was approximately (i.e. within the limitations of the algorithms) distributed as a uniform distribution with values between zero and one. The inverse cumulative distribution function was applied to each psuedo-random number to calculate a simulated time-to-failure. The accuracy of the Monte Carlo method generally improves with increasing numbers of simulation sets analyzed. Here, the term "simulation set" refers to the process of generating "N" simulated random samples and then calculating sample statistics of interest from those "N" samples. Concerning the numbers of simulation sets required, the advice found in the open literature was not consistent, although 5,000 simulation sets is often considered as adequate. After some study, it was decided that for the present work 20,000 simulation sets were sufficient to characterize the sampling distributions. Evidence concerning the adequacy of 20,000 simulation sets is provided in the text to follow.

Validation of Monte Carlo Scheme.—To validate the approach and computer code written to simulate random sampling from a Weibull population, a validation study was completed. The validation study consisted of five steps:

1. define a distribution representing a population of times-to-failures;
2. simulate the random selection of "N" samples from the population;
3. calculate sample statistics of interest from the "N" samples;
4. repeat steps one through three to complete 20,000 simulation sets;
5. analyze the collection of 20,000 sample statistics to characterize the sampling distribution.

Two populations representing times-to-failure were defined. One population was defined as a Weibull distribution, and the second population was defined as a normal distribution. The normal distribution was included in the validation study because the sampling distribution for the sample mean is defined by an exact analytical expression. Distribution parameters were selected so that the two populations would be similar (fig. 3.3.1). With no loss of generality, the scale and threshold parameters of the Weibull distribution were selected to be equal to 1.0 and 0.0, respectively. Making use of the relations provided by Cohen (ref. 3.23), the shape factor that allows for the mode and the mean of the distribution to coincide (the value 3.312) was found. The mode and mean of such a distribution have the value 0.8972 while the median of the distribution has the value 0.8953. The mode of the normal distribution was selected to match that of the Weibull distribution. Random sampling from these two distributions were simulated using the multiplicative congruential method for psuedo-random sampling and the inverse cumulative distribution method to calculate a simulated time-to-failure. Studies were done for sample sizes of 10 and 30.

The sample statistic of interest for the validation study was selected to be the 50-percent life, or median value. For the case of selecting samples from the normal distribution, the sample 50-percent life was calculated as the sample mean. For the case of selecting samples from the Weibull distribution, the sample 50-percent life was calculated making use of the best-fit distribution parameters. These best-fit distribution parameters were found by three methods, namely two-parameter least-squares regression of exact median ranks, three-parameter least-squares regression of exact median ranks, and two-parameter maximum likelihood. Details of the three methods are described in chapter 2 and section 3.2.

To characterize the sampling distributions, the set of 20,000 estimated 50-percent lives, calculated as described in the previous paragraph, was collected, and the results were sorted. From the sorted estimates, the 10th percentile, median, and 90th percentile were determined. Results of the validation study are provided in Table 3.3.1. Focusing for the moment on the results for the normal distribution, the exact

theory and Monte Carlo solutions are very close, with less than 0.5 percent deviation. The results provide evidence that the selected psuedo-random number generator is appropriate. The results also indicate that 20,000 simulation sets are probably sufficient to characterize the sampling distribution. Next, focusing attention on the results for the Weibull distribution, it is expected that the sampling distribution for the particular Weibull distribution of figure 3.3.1 should be similar to that for the normal distribution. Indeed, the location and breadth of the sampling distributions are similar. The deviation of the results for the Weibull distribution relative to the exact theory for the normal distribution are due to the slight differences in the probability distribution functions and due to the properties of the estimating methods. The results of this validation study demonstrate that the Monte Carlo scheme produces results with reasonable engineering accuracy. It was noted that the breadth of the sampling distribution for the case of 3-parameter least-squares regression is somewhat larger than the other 2-parameter based methods, providing a first clue that relaxing the assumption of a known threshold parameter will come with the consequence of larger confidence intervals. The Monte Carlo scheme was substantially validated by the results just presented. Further evidence concerning the adequacy of the selected psuedo-random number generator and the sufficiency of 20,000 simulation sets are provided in the text that follows immediately.

Required Number of Monte Carlo Simulation Sets.—To provide confirming evidence that 20,000 simulation sets will sufficiently characterize the sampling distributions of interest, a study was devised and completed. Monte Carlo simulation was used to simulate the random sampling from a Weibull distribution having a scale parameter equal to 1.0, a shape parameter equal to 1.2, and a threshold parameter equal to 0.0. The shape parameter value was selected as one typical for gear fatigue data. A simulation set consisted of generating ten psuedo-random samples from the Weibull distribution, estimating the parameters of the distribution from the 10 samples, and calculating the 10- and 50-percent life estimates using the parameter estimates. Parameter estimates were calculated using both 2-parameter and 3-parameter least-squares regression. The parameter and life estimates were collected and sorted. From the sorted estimates, the 5th, 50th, and 95th percentiles were plotted as a function of the number of simulations sets completed. These percentiles were chosen to provide a visual assessment of the location and breadth of the sampling distributions. All calculations were completed twice using the two different pseudo-random number generator schemes as described in the previous text.

Results of the Monte Carlo simulations are provided in figures 3.3.2 to 3.3.6. The plots on these figures show that the sampling distributions can be reasonably established using 20,000 simulation sets. The plots of the parameter estimates (figs. 3.3.2 to 3.3.4) illustrate that the sampling distributions are not symmetric, especially for the shape and threshold parameters. Even with 20,000 simulation sets, the extreme tails have not been established with precision. However, the parameter estimates are not the goal but a means to an end. The final goal of the analysis is to provide estimates of percentiles of interest. In this work, the 10- and 50-percentiles of the distribution describing fatigue life have been chosen to be the ones of interest. Plots of the life estimates are provided in figures 3.3.5 and 3.3.6. The sampling distributions have essentially been established using 20,000 simulation sets, and little is gained by extending the number of simulations. Both random number generators used produced very similar results, again confirming that either one would be appropriate for further studies. The slight differences that arose from using the two different random number generators is attributed to making use of two differing strings of (pseudo)random numbers. Figure 3.3.5 showing the 10-percent life estimates illustrates that the 3-parameter regression method produces a sampling distribution with somewhat greater breadth as compared to the 2-parameter regression method.

As a final check that a single set of 20,000 simulation sets is sufficient for purposes of establishing the sampling distribution for 10- and 50-percentile life estimates, the process was repeated 60 times resulting in a total of 1.2 million simulation sets. The calculations were done using consecutive seeds for the random number generators, that is, a string of 12 million random numbers was broken up into 1.2 million sets of 10 numbers each. Results of the study are provided in table 3.3.2. The ranges of values are significantly less than the breadths of the sampling distributions. The data of table 3.3.1 validate that either of the two random number generators would be appropriate for further studies. From this point forward, all Monte Carlo simulation studies were completed using the multiplicative congruential method

(refs. 3.20 and 3.21) for the simulation of random sampling, and a group of 20,000 simulations sets were used to establish, with reasonable engineering accuracy, the properties of a sampling distribution.

3.4 Three Methods for Estimating Weibull Sample Statistics—Qualitative Assessment

A study was completed to make a qualitative assessment of three methods for estimating Weibull distribution parameters. The three methods included in the study were the 2-parameter regression, 3-parameter regression, and 2-parameter maximum likelihood methods. The regression methods were implemented using the exact median rank method and the least-squares fitting criteria. Section 3.2 describes additional details about the implementation of the three methods.

The qualitative assessment of the three methods was done using the Monte Carlo method (as described in section 3.3) to simulate random sampling. The random sampling was from a Weibull distribution defined by a scale parameter equal to 1.0, a shape parameter equal to 2.0, and a threshold parameter equal to 0.0. A simulation set consisted of the process of simulating “N” random samples, fitting the distribution parameters to the group of “N” random samples, and collecting the resulting predicted parameters and percentiles. To create approximate sampling distributions, 20,000 simulation sets were completed, and the results are displayed as histograms. For all studies, a single, randomly selected starting seed value was used for the random number generator. Therefore, for each parameter fitting method studied, the same string of pseudo-random numbers was employed. In this study, there was no attempt to simulate a censoring scheme. All datasets were analyzed as complete data. (The influence of censoring was considered as a separate study, and the results are reported in section 3.8.)

Histograms representing approximate sampling distributions are provided in figures 3.4.1 to 3.4.11. The histograms depict the results for samples sizes of both 10 and 30, thereby covering the usual range of samples sizes for many gear fatigue studies. When appropriate, the true value of the parameter or percentile is labeled on the graph abscissa, the true value equaling the value for the population from which random samples were generated. A qualitative assessment of the results follows.

Approximate samplings distributions for the shape parameter are provided in figure 3.4.1 for sample size of 10 and in figure 3.4.2 for sample size of 30. As expected, the sampling distribution for the 3-parameter method has a somewhat greater breadth as compared to the distributions for the 2-parameter methods. Furthermore, the 3-parameter method appears to be less accurate in the sense that the mode of the sampling distribution does not equal the true value as well as the 2-parameter methods. The maximum likelihood method offers some advantage in precision (less breadth of the sampling distribution) for the larger sample size of 30 (fig. 3.4.2(c)). For estimating the shape parameter, the maximum likelihood method appears to be the method of choice for the case of sample sizes in the range of 10 to 30, a population with a zero-valued threshold parameter, and no censoring.

Approximate samplings distributions for the scale parameter are provided in figure 3.4.3 for sample size of 10 and in figure 3.4.4 for sample size of 30. The accuracy and precision of the 2-parameter methods seem to be roughly equivalent for both sample sizes. The scale parameter estimates provided by the 3-parameter method are relatively broad, and for the smaller sample size of 10 the mode of the sampling distribution is significantly smaller than the true value of 1.0. For estimating the scale parameter, these data alone show no difference between the 2-parameter regression and the maximum likelihood methods.

Approximate sampling distributions for the threshold parameter estimates are provided in figure 3.4.5. For both sample sizes examined, the resulting sampling distributions are skewed left. Note that the position of the mode of the sampling distribution changes relative to the true value of the parameter as a function of the number of samples. The data show that for small sample sizes, the left-side tail of the distribution is quite long. This reinforces the observation made earlier in this text concerning figure 3.2.2, that is, in cases when a negative threshold value produces the minimal sum of squared-errors, the sum of squared-errors is fairly insensitive to neighboring assumed threshold values.

Approximate sampling distributions for the 10-percent life predictions are provided in figure 3.4.6 for a sample size of 10 and figure 3.4.7 for a sample size of 30. Recall that for the shape and scale parameter

estimates, the sampling distributions for the 2-parameter methods were distinctively different from the 3-parameter method. In contrast, the sampling distributions for the 10-percent life predictions are remarkably similar. However, careful examination of figure 3.4.6 shows that for small sample sizes, the mode of the sampling distribution for the maximum likelihood method is more closely aligned to the true value than are the modes of the sampling distributions for the regression-based methods.

Approximate sampling distributions for the 50-percent life predictions are provided in figure 3.4.8 for a sample size of 10 and figure 3.4.9 for a sample size of 30. For the conditions studied here, there is little difference among the three methods for estimating the 50-percent life. The sampling distributions are nearly symmetric, and the shape of the distribution approximates a normal distribution.

The coefficient of determination, often called the r-squared statistic of the regression, is often used as a measure of the goodness-of-fit for regression analysis. The approximate sampling distributions for the coefficient of determination are provided in figure 3.4.10 for a sample size of 10 and figure 3.4.11 for a sample size of 30. These figures would seem to indicate that the 3-parameter regression method is preferred over the 2-parameter regression method. However, the sampling distributions for the life and parameter estimates indicate that at least for the conditions of this study, the 2-parameter method is as good as or better than the 3-parameter method. Therefore, for a particular dataset, one should not rely on the coefficient of determination as the means for selecting one regression-based method over another. However, approximate sampling distributions determined by Monte Carlo simulation, such as provided in figures 3.4.10 and 3.4.11, can be useful for checking whether or not one has presumed an appropriate form for the population distribution. For example, if the coefficient of determination has a value located far in a tail of the sampling distribution, then one may want to reconsider whether the population is distributed as a Weibull distribution.

In this section, Monte Carlo simulation was used to produce approximate sampling distributions of statistics of interest. The sampling distributions were presented in graphical form as histograms, and a qualitative assessment was completed. In general, the 2-parameter maximum likelihood method performed as well or better than the regression-based methods. This preceding conclusion is based on a qualitative assessment of the precision and accuracy of the estimates. However, one is cautioned that this assessment was limited to the case of a Weibull population with a shape factor of 2.0, a zero-valued threshold parameter, and complete data (no censoring). To provide more guidance for a final recommendation, quantitative assessments were completed, and the results are presented in later sections of this chapter. The next section describes a methodology for quantitative assessments that will complement the qualitative assessment just presented.

3.5 A Methodology for Quantitative Assessment of Statistical Procedures

In the preceding two sections, a Monte Carlo method was first validated and then used to provide a qualitative assessment of three methods for estimating a Weibull distribution from sample data. The scope of the qualitative assessment was limited, and it was desired to complete further quantitative assessments to provide a more complete evaluation. In this section, a quantitative methodology is defined and illustrated.

Precision and Accuracy of Sample Statistics.—To assess the performance of a statistical procedure, one needs to consider both the accuracy and precision of the sample statistics. The accuracy concerns how well the central tendency of the sampling distribution reflects the true value of the population. The precision concerns the breadth of the sampling distribution. In the present study, approximate sampling distributions are determined using Monte Carlo methods. The text to follow provides definitions and methods used in the present work to evaluate the approximate sampling distributions.

The accuracy of a sample statistic is stated in terms of the central tendency of the sampling distribution. There exist three popular measures of central tendency of a distribution, namely the expected value, the median, and the mode. Traditionally, the accuracy of a sample statistic has been stated in terms of bias, defined as the expected value of the sampling distribution minus the true value of the statistic. For some sample statistics, such as the sample mean of random samples from a normal distribution, the

sampling distribution will be symmetric. In such cases, the expected value, median, and mode coincide, and so it is natural to seek out statistical procedures that have zero bias. In the case of the Weibull distribution, the sampling distributions of statistics of interest are not necessarily symmetric (see, for example, fig. 3.4.1). Some have considered a biased estimator to be undesired even for the cases of asymmetric sampling distributions, and work has been done to develop unbiased estimators for Weibull statistics (refs. 3.10, 3.12, 3.13, 3.15, 3.17, and 3.24). However, as discussed by Cacciari, et al. (ref. 3.13), depending on one's application the unbiasing methods may be inappropriate, and such methods might provide misleading results. In the present work, the accuracy of a sample statistic is not viewed in terms of the traditionally defined bias but instead by a newly introduced term that has been named the "mode-based bias". The mode-based bias is defined as the value of the mode of the sampling distribution minus the true value of the statistic. The concept of the "mode-based bias" will now be discussed with the aid of figure 3.5.1. The figure depicts approximate sampling distributions created in the manner of those of figures 3.4.1 to 3.4.10 except the histograms have been replaced by a smooth curve fit to the histogram values. Note that for the case of the 2-parameter regression method (fig. 3.5.1(a) and (c)), the mode of the sampling distribution does not equal the true value, and the relative positions of the mode to the true value seems to depend on the sample size. On the other hand, for the case of the maximum likelihood method (fig. 3.5.1(b) and (d)), the mode of the sampling distribution closely matches the true value, and the relative position of the mode to the true value seems to be relatively insensitive to the sample size. In this work, the mode was selected as the preferred measure of central tendency, and so the mode-based bias was adopted as the measure of the accuracy of sample statistics.

To quantify the mode-based bias, the following method was adopted for determining the mode of a sampling distribution. An array of 20,000 sample statistics was generated using Monte Carlo simulation (section 3.3). A histogram was created from the 20,000 estimates using an interval size to produce approximately 2000 intervals over the range of the data. A locally-weighted regression method (known as lowess, ref. 3.25) was then applied to the values of the histogram, acting as a curve-smoothing technique. The locally-weighted regression smoothing factor used was 0.10. The histogram and locally smoothed curve were plotted for a visual check, and the abscissa value resulting in the maximum value of the locally smooth curve was found to provide the location of the mode of the approximate sampling distribution. An example result is provided in figure 3.5.2. The figure shows that the locally smooth curve is a reasonable representation of the histogram data, and the location of the mode is marked on the abscissa of figure 3.5.2(b). To check for the robustness of the just described method, the dataset used to produce figure 3.5.2 was analyzed again using ranges of values for the histogram interval size and for the locally-weighted regression smoothing factor. The location of the mode did not change (to three significant figures) for histogram interval sizes comprising [0.02, 0.01, 0.005, 0.001, and 0.0005] while employing a smoothing factor of 0.10. Likewise, the location of the mode did not change (to three significant figures) for smoothing factors comprising [0.30, 0.20, 0.10, 0.05, and 0.02] while employing a histogram interval size of 0.001. These calculations demonstrate that the proposed method for determining the mode of the sampling distribution is very robust with respect to the factors selected for histogram interval size and curve smoothing.

To quantify the precision of the estimates, the "middle" 90-percent of the sampling distribution was selected by the definition of the distribution's breadth. The "middle" 90-percent was determined by first sorting the 20,000 sample statistic estimates resulting from the Monte Carlo simulation. Then, the estimates representing the 5th and 95th percentiles were selected from the sorted list. These estimates defined the endpoints of the interval containing 90-percent of the approximate sampling distribution. An example result is provided by the symbols located on the abscissa of figure 3.5.2(b).

A methodology has been developed to evaluate the accuracy and precision of statistical procedures. Approximate sampling distributions are created making use of 20,000 Monte Carlo simulation sets. The results of the Monte Carlo simulations are evaluated to determine the central tendency and breadth of the sampling distribution. The accuracy is quantified as the value of the mode-based bias. The precision is quantified by the interval bound by the 5th and 95th percentiles of the ordered estimates.

3.6 Presumption of a Zero-valued Threshold Parameter—A Critical Review

Background and Motivation.—In the present work, the Weibull distribution has been selected as a parametric model for the distribution of gear fatigue life. Traditionally, when the Weibull distribution has been used to model gear fatigue life distributions, the threshold parameter has been presumed to equal zero (that is, a two-parameter distribution has been presumed). Usually, no justification is given for the presumption of a zero-valued threshold parameter. At the time of the introduction of the Weibull distribution, the computing power required to estimate all three parameters from sample data would have been considered a significant resource. It is plausible that the computing cost and analyst's time required to make the necessary calculations to estimate all three parameters were, at one time, a compelling reason to employ the presumption of a zero-valued threshold. Today, appropriate computing power and numerical methods are readily available to estimate all three parameters of the distribution. The preceding statements provide one motivation for making a critical review of the presumption of a zero-valued threshold parameter.

A second motivation for considering the idea of a non-zero valued threshold parameter comes from knowledge of failure processes. Considering the physics of failure mechanisms, one could argue that the true threshold value for gear surface-fatigue failure life distributions cannot be identically equal to zero. For such a presumption, there exists a small, non-zero probability of fatigue failure during the first load cycle. However, a failure during the first load cycle would not be considered a fatigue failure but instead be considered as failure by another mechanism. Furthermore, for the loads applied in typical gear fatigue testing, fatigue failures have occurred only after millions of stress cycles. Given these ideas and facts, a critical review of the presumption of a zero-valued threshold was included in the present work.

In the preceding paragraph, an argument was set forth that the threshold parameter cannot be identically equal to zero. Still, the presumption of a zero-valued threshold may be appropriate if the zero-valued presumption is a good approximation to the true value. This was demonstrated the results of the qualitative assessment of the statistical methods for estimating sample statistics (section 3.4). It was shown that if the true threshold parameter was identically equal to zero, then the 2-parameter methods were better performers than was the 3-parameter method. It would also be anticipated that the 2-parameter methods to be the better performers if the true value for the threshold parameter was non-zero but “small enough”. On the other hand, it would be anticipated that the 3-parameter method to be the better performer if the true value for the threshold parameter was “large enough”. With these ideas in mind, the critical review to be conducted was done with the view that although the true value of the threshold parameter is not identically zero, the zero-valued presumption may indeed be reasonable and may indeed prove to be advantageous.

The review of the presumption of a zero-valued threshold value was done not to support any theoretical arguments but instead with a practical goal in mind. In the end, the goal is to provide estimates of the 10-percent life of a population of gears, and this estimate will come about from a limited number of tests of randomly selected samples from the population. To achieve this goal, two complementary studies were completed. The first study was done to assess the precision and accuracy of 10-percent life estimates as a function of the true value of the threshold parameter. Thereby, the study quantifies the implications of approximating a non-zero threshold as being identically equal to zero. The second study was done using data from 12 gear fatigue experiments to assess, in a qualitative manner, a likely range for the true value of the threshold parameter. Based on the evidence from these two studies, a final recommendation will be provided.

Influence of the true value of the threshold parameter.—An evaluation was done to assess the influence of the true value of the threshold parameter on the accuracy and precision of 10-percent life estimates. The quantitative evaluation was done using the methods of section 3.5. The 10-percent life estimates were made using three methods, namely the 2-parameter least-squares regression of exact median ranks, the 3-parameter least-squares regression of exact median ranks, and the 2-parameter maximum likelihood method.

To make the evaluation, four Weibull distributions were defined with differing threshold values. The four distributions were defined such that the middle parts of the distributions would have similar lives. The “baseline” distribution was defined to have a shape parameter equal to 2.0, a scale parameter equal to 1.0, and a threshold parameter equal to 0.0. Next, three distributions were defined with respective threshold values equal to 0.1, 0.2 and 0.3 while also having “similar” lives to that of the baseline distribution. The shape and scale parameters for the three distributions having positive valued threshold parameters were selected to minimize, in the least-squares sense, the deviations of the lives for the following set of values for cumulative percent failed: [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9]. Plots of the four similar Weibull distributions with differing threshold values are provided in figure 3.6.1, and the parameters defining these distributions are provided in Table 3.6.1. The life corresponding to the 63rd percentile is sometimes called the characteristic life. All four distributions have characteristic lives of approximately 1.0.

The precision and accuracy of the 10-percent life estimates were determined using the methods of section 3.5. Evaluations were made for the cases of sample sizes of 10 and 30 and for the condition of no censoring. The results are provided in figures 3.6.2 and 3.6.3.

First, consider the results for 10 samples (fig. 3.6.2). For a true threshold value of 0.10 or less, the maximum likelihood method is the best performer. The mode-based bias is less than that for the regression-based methods, and the precision of the estimates are also better or comparable. For a true threshold value equal to 0.20, the picture is somewhat mixed in that the 2-parameter regression method has the smallest interval size but the largest bias while the maximum likelihood method has the smallest bias but the largest interval size. For a true value equal to 0.30, the 3-parameter regression method is clearly the best performer since it provides both the smallest interval size and smallest bias. One might consider the threshold value of 0.20 as an approximate “break-even” point. For true threshold values less than 0.20, one would tend to choose the maximum likelihood method as the best performer. On the other hand, for true threshold values greater than 0.20 one would tend to choose the 3-parameter regression method as the best performer.

Next, consider the results for 30 samples (fig. 3.6.3). In general, the trends are similar to those, as just discussed, for 10 samples. For small true values of the threshold parameter, the maximum likelihood method is the best performer. For relatively large values of the threshold parameter, the 3-parameter regression method is the best performer. One might consider the threshold value of 0.15 as an approximate “break-even” point since at this value the bias of the maximum likelihood and 3-parameter regression-based estimates are about equal and the interval sizes comparable.

As was expected, for threshold values that are “small enough,” the zero-valued presumption is a good approximation and therefore advantageous. The assessment indicates that for characteristic lives of about 1.0, threshold values in the range [0.0~0.15] can be considered as “small enough” to justify the zero-valued threshold presumption. Likewise, for threshold values that are “large enough,” the zero-valued presumption can be an inadequate approximation, and in such a case it is advantageous to estimate the threshold parameter from the sample data. The assessment indicated that for characteristic lives of about 1.0, threshold values greater than about 0.2 can be considered as “large enough” such that one may prefer to estimate the threshold value from the sample data rather than presume a value of 0.0. The next step of this evaluation was to complete a qualitative study of typical gear fatigue datasets to seek out evidence as to the likely true values of threshold parameters (relative to the characteristic lives of the datasets).

Qualitative assessment of likely values for the threshold parameter.—To provide a qualitative evaluation of the likely values of the threshold parameters (relative to the characteristic lives of the datasets), experimental and simulated test data were plotted using Weibull coordinates. The concept for the evaluation is illustrated per figure 3.6.4. The four Weibull distributions that were depicted in figure 3.6.1 using linear coordinates are provided again but, in this instance (fig. 3.6.4) the data are plotted using Weibull coordinates. One can see that for the case of a zero-valued threshold parameter (fig. 3.6.4(a)), the distribution plots as a straight line. However, as the relative value of the threshold parameter becomes larger, the lower-left portion of the curve deviates from a straight line, and the degree of curvature increases as the threshold value increases (fig. 3.6.4(b) to (d)).

To provide experimental data for evaluation, data from 12 gear fatigue experiments representing typical datasets were gathered (refs. 3.26 to 3.34). For these 12 datasets, the number of completed tests ranged from 17 to 21 with the most common value being 20. The data of these references were published in graphic form. To obtain the numerical values, the plots were digitized to obtain the times-to-failure. To prepare for a qualitative assessment, the experimental data were plotted, using Weibull coordinates, at the exact-median-rank plotting positions (ref. 3.4). The datasets included a limited number of censored data points. The censored tests were considered to determine the exact-median-rank plotting positions, but the censored data points were not plotted. To provide for a common scale on the abscissa, the data were normalized such that the data point to be plotted closest to the 63rd cumulative failed position on the ordinate was plotted as a life of 1.0, and the remaining data were rescaled accordingly.

Next, to provide sets of simulated data for comparison, the Monte Carlo method was used to simulate the random sampling effect. The simulated random sampling was done from the four Weibull distributions used for qualitative assessments of the preceding paragraphs (fig. 3.6.1 and Table 3.6.1), and twelve datasets of 20 pseudo-randomly selected data points each were provided from each distribution.

Plots of the 12 sets of gear fatigue experimental data that were gathered and rescaled are provided in figure 3.6.5. Plots of data from four Weibull distributions with simulated random sampling effects are provided in figures 3.6.6 to 3.6.9. Whether experimental or simulated data, the data points will not form perfect straight lines for two reasons. One reason why the data deviate from a straight line is the random sampling effect, and the second (potential) reason for the deviation from a straight line could be the existence of a positive valued threshold parameter. The data from the Monte Carlo simulations (figs. 3.6.6 to 3.6.9) provide a calibration of one's expectation of the "look" of the plotted data having the superimposed effects of random sampling and (potential) non-zero valued threshold parameters.

Comparing first the experimental data (fig. 3.6.5) to the simulated data for the case of a zero-valued threshold parameter (fig. 3.6.6), it is apparent that the experimental data has, in almost every case, the appearance of a trend with a least some curvature in the lower-left part of the plot. On the other hand, the simulated data appear, relatively speaking, as fully straight lines. Therefore, the experimental evidence does not support the presumption of a zero-valued threshold parameter.

Consider now the case of a threshold value equaling 0.3, the largest of the assessed threshold values (fig. 3.6.9). It is apparent that, in almost every case, the curvature extends for nearly the entire range of the data. It is also apparent that, in almost every case, the data suggests a clearly defined asymptote as the lower portion of the data tends toward the appearance of a vertical line. In contrast, the experimental data (fig. 3.6.5) do not, in all cases, include the appearance of curvature for the entire dataset nor, in all cases, the appearance of a clearly defined vertical asymptote. Therefore, the experimental data do not support the presumption of a threshold parameter with a value (relative to a characteristic life of 1.0) as large as 0.3.

Next consider the cases for threshold values of 0.1 and 0.2 (figs. 3.6.7 and 3.6.8). Comparing these plots to those of the experimental data (fig. 3.6.5), it is apparent that some of the experimental datasets have a form with relatively little curvature and thereby more resemble the forms of the simulated data having a threshold value of 0.1 (fig. 3.6.7). On the other hand, some of the experimental datasets have relatively more curvature and thereby more resemble the forms for the simulated data having a threshold value of 0.2 (fig. 3.6.8). Overall, the experimental evidence suggests that the most likely values for the threshold parameters are within the range of [0.1~0.2] rather than more extreme values examined (0.0 or 0.3).

Implications of the quantitative and qualitative evaluations.—In the preceding paragraphs, quantitative and qualitative assessments were completed to provide guidance as to whether or not the traditional presumption of a zero-valued threshold parameter is an adequate approximation. The results of the studies were assessed keeping in mind that the end goal is to provide estimates of the 10-percent lives of gear fatigue life distributions using sample data. From the quantitative study of the influence of the true threshold value on the performance of the estimating methods, one can conclude that the 2-parameter estimating methods are robust to "small" positive values for the true threshold parameter. The accuracy and precision of the estimates were not significantly effected unless the value of the threshold parameter

was larger than about 0.2 (relative to a characteristic life of 1.0). The qualitative assessment of 12 sets of experimental data indicated that the most likely values for the threshold parameters (relative to characteristic lives of 1.0) were within the range [0.1~0.2]. Therefore, although the experimental evidence suggests that the true values for the threshold parameters are not identically zero, the most likely values are small enough such that the zero-valued presumption is a good approximation. These findings for gear fatigue data are in agreement with the findings of a study of bearing fatigue data. Tallian (ref. 3.35) developed and applied a normalizing scheme to create a single dataset of 2300 fatigue tests for bearings, and from that dataset assessed the goodness of fit to a 2-parameter Weibull distribution. Although he discovered that the data deviated from the Weibull distribution in the extreme tails, the likely value of the threshold was a small fraction of the characteristic life and, therefore, the 2-parameter Weibull distribution provided a good fit for purposes of estimating the 10-percent life.

Considering the results of this study of gear fatigue data and the results from analysis of bearing data, a final recommendation can be provided. For the purpose of estimating 10-percent lives from sample data, it is recommended that gear fatigue data should be modeled as having a zero-valued threshold parameter. With this recommendation in mind, the 3-parameter regression method was not included as part of the studies of the influence of sample size and censoring (to be presented in the next two sections).

3.7 Influence of the Number of Samples on the Accuracy and Precision of 10-Percent Life Estimates

The influence of the number of samples in a dataset on the accuracy and precision of 10-percent life estimates was studied. Approximate sampling distributions were created using the Monte Carlo method described in section 3.3. The accuracy and precision was assessed by calculating the properties of the approximate sampling distributions per the methods of section 3.5. The assessments were done for sample sizes ranging from 7 to 100 and for two separate Weibull distributions with true values for the shape parameter of 1.0 and 2.0. The scale parameters of the two distributions were selected to have equal 10-percent lives. This study was limited to cases of complete data (that is, all samples were treated as having failed). The influence of suspending tests at prespecified times (censoring) was studied separately, and those results will be reported in section 3.8.

Approximate sampling distributions for 10-percent life estimates for the case of selecting samples from a distribution with a shape parameter value of 2.0 are illustrated in figure 3.7.1. The properties of the exact median rank regression method are provided in figure 3.7.1(a), while those of the maximum likelihood method are provided in figure 3.7.1(b). The properties of the two distributions are similar. However, the mode of the maximum likelihood method better approximates the true value than does the mode of the exact median rank regression method.

The properties of the approximate sampling distributions for 10-percent life estimates for the case of selecting samples from a distribution with a shape parameter value of 1.0 are illustrated in figure 3.7.2. The properties of the exact median rank regression method are provided in figure 3.7.2(a), while those of the maximum likelihood method are provided in figure 3.7.2(b). The properties of the two distributions are similar. However, the mode of the maximum likelihood method better approximates the true value than does the mode of the exact median rank regression method, especially for sample sizes less than 40. Pointing out another difference, the value locating the 95th percentile of the ordered estimates is somewhat larger for the maximum likelihood method for sample sizes less than about 15.

The accuracy and precision of the 10-percent life estimates are depicted graphically in figures 3.7.3 and 3.7.4. For the case of random samples drawn from a population with shape parameter of 2.0 (fig. 3.7.3), the maximum likelihood method is clearly the better performing method. For sample sizes greater than about 12, both the accuracy and precision of the maximum likelihood method is better. For sample sizes less than about 12, the precision is slightly better for the exact median rank regression method, but the maximum likelihood method provides more accurate estimates. Practically speaking, for the entire range of sample sizes studied the maximum likelihood method performed as well or better than the regression-based method for the case of a true shape factor of 2.0.

The accuracy and precision of the 10-percent life estimates for the case of random samples drawn from a population with shape parameter of 1.0 are depicted graphically in figure 3.7.4. For sample sizes larger than about 30, the maximum likelihood method provides both better accuracy and better precision. For sample sizes less than about 30, the overall relative performance is somewhat mixed. The maximum likelihood method provides better accuracy, but the regression-based method provides better precision. Considering the overall performance, there is no compelling reason to select one method over the other for samples sizes less than about 30, while the maximum likelihood method is the better performer for sample sizes greater than about 30.

In this section, the accuracy and precision of 10-percent life estimates were studied as a function of the number of samples. The study described in this section was limited to the conditions of complete data, that is to cases having no censoring. For all cases studied, the accuracy was better for the maximum likelihood method. For the case of a true shape factor of 2.0, the maximum likelihood method also had a better precision. In contrast, for small numbers of samples and a true shape factor of 1.0, the precision of the maximum likelihood was not as good as the regression-based method. Still, for such a condition the accuracy was better, and for practical purposes the overall performance can be considered roughly equivalent. This study suggests that, in general, the maximum likelihood method is the method of choice.

3.8 Influence of Censoring on the Accuracy and Precision of 10-Percent Life Estimates

When experimentalists evaluate high cycle fatigue life, the test procedure often makes use of censoring. That is, some tests are suspended even though no fatigue has occurred. Effective use of censoring can allow for reaching a conclusion with less total test time as compared to conducting all tests to failure. In past work, gear fatigue testing has made use of censoring, and it is anticipated that censoring will be used in future work. Therefore, a study was completed to assess the influence of censoring on the accuracy and precision of 10-percent life estimates.

To study the influence of censoring on the performance of the statistical methods, Monte Carlo simulation was used to simulate random sampling from a Weibull distribution with a scale parameter equal to 1.0, a shape parameter equal to 2.0, and a threshold parameter equal to 0.0. Such a distribution has a 10-percent life equal to 0.3246. Censoring was included in the study as follows. A censoring time was selected, and if the simulated time-to-failure exceeded the censoring time, then that particular simulated fatigue test result was treated as a test suspended at the defined censoring time with no failure. However, if treating a particular simulation set in such a manner would result in only one or zero failures, then the two smallest simulated times to failure were considered as completed tests. Censoring times ranging from 3.246 to 0.4869 were studied, and so the ratio of the censoring time to the true 10-percent life ranged from 10 to 1.5. Studies were done for sample sizes of 15 and 30. The approximate sampling distributions produced by the Monte Carlo simulations were evaluated for accuracy and precision using the methods described in section 3.5.

Analysis of the Monte Carlo simulation results showed that the accuracy of the 10-percent life estimates (as measured by the mode-based bias) did not change over the range of censoring times studied. However, the precision of the estimates (as measured by the breadth of the intervals as bound by the 5th and 95th percentiles of the ordered estimates) was influenced by the censoring time. The relationship of the interval size to the censoring time is illustrated in figure 3.8.1. In this figure, the interval sizes and censoring times are provided as ratios relative to the true value for the 10-percent life. For the larger sample size of 30, the interval size is smaller for the maximum likelihood method relative to the regression-based method for all censoring times studied. The trend for the interval size as censoring times are reduced is nearly linear for all but the smallest censoring times that nearly approach the true 10-percent life. For the smaller sample size of 15, the interval size is smaller for the maximum likelihood method relative to the regression method for most of the range of censoring times studied. However, for the case of censoring times less than about four times the true 10-percent life and 15 samples, the interval size for the maximum likelihood method is greater than that for the regression method. In general, the

maximum likelihood method performed better than the regression method except for the case of a small values for the censoring time and few total samples tested.

To provide more insight concerning the behavior of the maximum likelihood method with respect to censored data, the results for the case of 30 samples were studied in more detail. The interval size, as was provided in figure 3.8.1, is provided again in figure 3.8.2 along with the mean value for the number of failures (fig. 3.8.2(a)) and the mean value of the relative total test time (fig. 3.8.2(b)). It is noted that the shapes of the curves for the mean value of the number of failures and the mean value of the total test time are similar. However, the reduction in the total test time is modest relative to the reduction in the number of failures. Comparing the smallest censoring time studied to the largest censoring time studied, the increase in interval size is relatively modest (about a 15 percent increase) while the reduction in the mean value for the total test time is significant (about a 50 percent reduction).

Additional insights concerning the behavior of the maximum likelihood method with respect to censoring were obtained by studying graphical representations of the simulation-based sampling distributions of 10-percent life estimates. Three such sampling distributions are provided in figure 3.8.3. The histograms are displayed as scatter plots, rather than the usual bar graphs, to provide clarity. The three cases depicted are the largest and two smallest censoring times studied. In general, the change in the sampling distribution as a function of the censoring time is slight, with a rather modest increase in the breadth of the distribution, no change in the location of the mode, and little change of the general shape. However, for the case of a censoring time equal to 0.49 (fig. 3.8.3(c)), there exists a “spike” in the histogram at the location equaling the censoring time. A careful inspection of the results for censoring time equal to 0.65 (fig. 3.8.3(b)) reveals that a modest “spike” exists, and its location is, likewise, equal to the censoring time.

Another view of the data in figure 3.8.3(c) is provided in figure 3.8.4 by plotting the estimated 10-percent life as a function of the proportion of tests completed. It is noted from this figure that the 10-percent life estimate will exceed the value 0.49 only for those simulation sets that contained 2 failures, and the estimate will equal approximately 0.49 for those simulation sets containing 3 failures. The number of such cases from the total of 20,000 simulation sets of the study are provided in the data of table 3.8.1. For the conditions studied, approximately 9-percent of the simulation sets contained 3 or fewer failures. It is noted that although there were 1228 simulation sets with 3 of 30 (or 0.10 proportion) failing (table 3.8.1), the range for these 1228 estimates was negligible (fig. 3.8.4). All of these data and observations seem to indicate that for the purpose of estimating the 10-percent life while employing the maximum likelihood method, the proportion of tests ending in failure should exceed one-tenth. The traits of the maximum likelihood method for the case of “heavy” censoring have been studied in a rigorous manner by Jeng and Meeker (ref. 3.18). They noted that with Type I censoring (the type studied here), the joint distribution of maximum likelihood estimators has a discrete component related to the random number of failures. They also noted poor behavior of simulation based confidence interval procedures when the quantile to be estimated is close to or equal to the proportion of tests resulting in failure. Such “exceptional” behavior leads to the recommendation that when employing Type I censoring, the 2-parameter maximum likelihood method should not be used to estimate the 10-percent life unless the proportion of tests failed exceeds one-tenth.

In general, this study of the influence of censoring on the performance of the maximum likelihood method and the regression of exact median rank method showed that the maximum likelihood method is usually the better performer. The accuracy and precision of the 10-percent life estimates is generally better except for certain “exceptional” cases. These “exceptional” cases are those that involve such an aggressive censoring time that the proportion of failures is close to, or less than, one-tenth. The results of this study of the influence of censoring leads to the recommendation of the maximum likelihood method over the 2-parameter regression of exact median ranks method.

As a final note on these studies of the influence of censoring, the work just described was done to assess the performance of the estimating methods. The method of study and results could also be viewed as a tool for optimizing test plans. Further studies using the Monte Carlo method as a planning tool for testing are beyond the scope of the present effort. However, such studies are recommended as future

extensions of this work since the results would likely be of great value to those conducting experiments to quantify high cycle fatigue lives.

3.9 A Summary of the Assessments of Statistical Methods and a Recommendation

In the preceding sections of this chapter, three methods for estimating sample statistics of a Weibull distribution were assessed and compared. In this section, the work and results are summarized, leading to a final conclusion and recommendation for the treatment of gear fatigue data.

In this work, it was assumed that the Weibull distribution is an appropriate one for description of gear fatigue data. Three methods for estimating the parameters and sample statistics of the Weibull distribution from randomly selected samples were selected for study, namely:

1. the 2-parameter least-squares regression method,
2. the 3-parameter least-squares regression method,
3. the 2-parameter maximum likelihood method.

These three methods were selected for study based on review of previous works (chapter 2). section 3.2 describes the details of the implementation of these three methods.

The primary tool used to assess and compare the methods was Monte Carlo simulation. The concept of the Monte Carlo simulation method along with the details of implementation for the present work are described in section 3.3. Section 3.3 also includes results of a study that validated the implemented Monte Carlo method. The result of a Monte Carlo simulation is a collection of sample statistics, each differing due the influence of the (simulated) random sampling process. To provide a qualitative assessment, the results can be displayed as a histogram thereby providing a visual assessment of the sampling distribution. To provide a quantitative assessment, a method is needed to quantify selected properties of the sampling distribution. A methodology for quantitative assessment of the accuracy and precision of a sample statistic is described in section 3.5. The accuracy of the sample statistic is based on the difference of the mode of the sampling distribution relative to the true value. The precision of the sample statistic is quantified in terms of the size of the interval containing the 5th through the 95th percentiles of the ordered estimates.

Qualitative assessments and comparisons of the three selected estimating methods are provided in section 3.4. The qualitative assessments were done for random sampling from a Weibull distribution defined by a scale parameter equal to 1.0, a shape parameter equal to 2.0, and a threshold parameter equal to 0.0. For this assessment, all random samples were considered as simulated data resulting in failure (no censoring). The assessments were done for two samples sizes of 10 and 30. In general, the 2-parameter maximum likelihood method performed as well or better than the regression-based methods.

Traditionally, when the Weibull distribution has been used for gear-fatigue-life data, the threshold parameter has been presumed to equal zero (that is, a two-parameter distribution has been presumed). The presumption of a zero-valued threshold parameter was critically examined and the results reported in section 3.6. Although the experimental evidence suggests that the true values for the threshold parameters are not identically zero, the most likely values are small enough such that the zero-valued presumption is a good approximation. Therefore, for the purpose of estimating 10-percent lives from sample data, it is recommended that gear fatigue data should be modeled as having a zero-valued threshold parameter. This recommendation for gear data matches the results of a study of bearing fatigue data (ref. 3.35) in that an assumed zero-valued threshold was deemed appropriate for the purpose of estimating the 10-percent lives of bearings.

The influence of the number of samples on the accuracy and precision of 10-percent life estimates was studied, and the results are provided in section 3.7. The assessments were done for sample sizes ranging from 7 to 100 and for two separate Weibull distributions with true values for the shape parameter of 1.0 and 2.0. This study of the influence of the number of samples was limited to cases of no censoring. For a true shape value of 2.0, the maximum likelihood method provided estimates that were both more

accurate and more precise. For a true shape value of 1.0 and small samples sizes, the maximum likelihood method was not as precise, but the method was more accurate and, therefore, for practical purposes it performs as well or better than the regression-based method.

The influence of censoring on the accuracy and precision of 10-percent life estimates was assessed, and the results are reported in section 3.8. In general, the maximum likelihood method was usually the better performer when applied to datasets that included censored data. The accuracy and precision of the 10-percent life estimates is generally better except for certain “exceptional” cases. These “exceptional” cases are those that involve such an aggressive censoring time that the proportion of failures is close to, or less than, one-tenth.

Considering all of the evidence provided by the assessments, the 2-parameter maximum likelihood method stands out as the method of choice over the regression-based methods for analysis of gear fatigue data. When random samples are taken from a Weibull distribution with a true value for the threshold parameter of zero, then the maximum likelihood method provides estimates of percentiles and parameters that are more accurate than the estimates provided by the regression-based methods. Furthermore, the maximum likelihood estimates are also often more precise than the regression-based estimates as measured by the breadth of the sampling distribution. For certain combinations of true parameter values and small sample sizes, the precision is better for the 2-parameter regression-based estimates while the accuracy is better for the maximum-likelihood-based estimates. In such situations, there is no compelling reason to select one method over the other. The maximum likelihood method is generally robust to censoring. For certain conditions of censoring the maximum likelihood method can produce poor behavior, but the conditions that produce such “exceptional” behavior are known and so can be avoided with proper test planning and execution.

3.10 Confidence Intervals—Implementation

In the previous sections of this chapter, concepts and methods for estimating sample statistics for gear fatigue life data were assessed and compared, and it was recommended to use the 2-parameter Weibull distribution and the maximum likelihood method for the statistical analysis and inference of gear fatigue data. However, it is best practice for researchers to report not only the estimates of the sample statistics but to also quantify the precision of those estimates by giving consideration to the random sampling effect. The most common way for one to quantify uncertainty due to the random sampling effect is to state a confidence interval. Confidence intervals are stated using some specified level of confidence. The level of confidence describes the performance of a confidence interval procedure and, thereby, expresses one’s confidence that a particular interval contains the quantity of interest. A brief overview of confidence interval methods was provided in section 2.4. Exact confidence interval methods for Weibull statistics are not available for datasets that include Type I censored data (ref. 3.36), and so approximate confidence interval methods must be employed. In this chapter, a confidence interval method that best complements the recommended estimating method (maximum likelihood) is selected and implemented.

Previous researchers have provided guidance for the selection of a method for confidence intervals for Weibull sample statistics (refs. 3.18 and 3.36). In these studies, the researchers compared the performance of two classes of confidence interval procedures. One class of procedures is based on certain sample statistics tending toward a normal distribution as the sample size increases. The second class of sample statistics is based on likelihood ratios and certain statistics that tend toward a chi-squared distribution as the sample size increases. The confidence intervals reported by most commercial software make use of the first of these classes, which are normal-approximate methods. However, both of the referenced studies found that the normal approximation was relatively poor for “small” samples sizes of less than about 50. It is noted that when gear surface fatigue life estimates based on experiments have been reported in the literature, the confidence interval for the estimate (if one is reported) is usually a normal-approximate based interval. The more recently developed methods based on likelihood ratios have been found to be better performers than the normal-approximate based methods (refs. 3.18 and 3.36), especially for the sample sizes commonly used for gear fatigue research. Jeng and Meeker (ref. 3.18) also

conducted extensive studies of a third class of confidence interval procedures that are based on parametric bootstrap principles. They consider that certain parametric bootstrap procedures offer some advantages over likelihood ratio based methods. However, such methods require extensive calculations and special software. For example, summarizing the work of Jeng and Meeker (ref. 3.18), the Parametric Bootstrap Signed Square-Root Log-Likelihood Ratio Procedure (PBSRLLR) was recommended as the best practice and “should be employed *when appropriate software becomes available*” (italics added for emphasis). Implementation of confidence interval methods based on parametric bootstrap concepts was considered beyond the scope of the present effort. Based on review of the existing literature and the current state of developments in the field of statistics, the likelihood ratio method for confidence intervals was selected for the present work.

Likelihood ratio based confidence interval procedures were implemented making use of the methods and examples provided by Meeker and Escobar (ref. 3.37). Using a two-parameter Weibull distribution as an example, the profile likelihood function for the shape parameter β is

$$R(\beta) = \max_{\eta} \left[\frac{L(\beta, \eta)}{L(\hat{\beta}, \hat{\eta})} \right]. \quad (3.10.1)$$

In the previous equation, for a fixed value of β , η is determined to provide the maximum value for the ratio. The denominator of the ratio is the likelihood value found using the maximum likelihood estimates of the parameters. Once the profile likelihood is determined, an approximate confidence interval for 100(1- α) percent confidence is given by

$$R(\beta) > \exp \left[-\chi^2_{(1-\alpha;1)} / 2 \right] \quad (3.10.2)$$

Because the maximum likelihood estimators have an invariance property, one can reparameterize the Weibull distribution function to be in terms of any percentile of interest, and the concept of the likelihood ratio still applies (ref. 3.37). Software was written and validated to implement the likelihood ratio based method for confidence intervals (ref. 3.9). The method will be applied to the data to be presented in chapter 4.

3.11 A New Method for Comparing Two Datasets to Assess Life Improvements

In the preceding sections of this chapter, concepts and methods for the statistical analysis of gear fatigue data, including confidence intervals, were studied and assessed, and best practices were recommended. However, often the ultimate goal of gear fatigue research is to determine whether or not two populations of gears have lives that are statistically significantly different. As such, a method is needed to compare two datasets that are representative of two populations having (potentially) differing fatigue life distributions.

Although much work has been published concerning the estimation of Weibull population parameters, surprisingly little work has been published on methods to make such a comparison. Johnson (ref. 3.38) provides a method for comparing two datasets on the basis of the ratios of sample mean lives or sample 10-percent lives. His method is based on normal-approximate theory. Jeng and Meeker (ref. 3.18) among others have demonstrated that normal-approximate theory performs poorly for small sample sizes and that normal-approximate theory provides nonconservative estimates of confidence. Meeker and Escobar (ref. 3.37) provide an example of comparing two datasets on the basis of the estimated 50-percent lives. However, in that example the lives are modeled as having normal distributions, and similar equations as used for the example are not available for the case of distributions modeled as Weibull distributions. McCool (ref. 3.39) developed a method for comparing populations via hypothesis testing. He makes use of certain sample statistics that are median unbiased and independent of the true Weibull

parameter values. His method for comparing populations on the basis of chosen quantiles requires an assumption that the two populations have a common but unknown shape parameter. The properties of his test statistics are established by Monte Carlo methods.

To enable the comparison of two sets of data on the basis of 10-percent lives (or any other chosen quantile) even for case that the true shape parameters differ, a new method is proposed. The method is based on the likelihood ratio and is a straightforward extension of the likelihood ratio based confidence interval methods. The method will be described in the paragraph to follow using the 10-percent lives as the basis for comparison, but the method is general and any quantile of interest could be examined.

The newly proposed method will now be described in a step-by-step manner. The first step of the newly proposed method is to form the null hypothesis that the 10-percent lives are equal. Since the lives are modeled as Weibull distributions, and each distribution has two free parameters, there exist an infinite number of solutions that permit the 10-percent lives of the two distributions to be equal. However, it is proposed that the common 10-percent life that maximizes the likelihood of the sample data is the value that is used to test the null hypothesis. Let the symbol θ denote a vector consisting of the four parameters that define the two Weibull populations,

$$\theta = (\beta_1, \eta_1, \beta_2, \eta_2). \quad (3.11.1)$$

Here, the subscripts denote populations (or datasets) one and two, respectively. Let θ_M denote the vector such that the four elements of the vector have the values representing the maximum likelihood solutions.

$$\theta_M = (\beta_{1M}, \eta_{1M}, \beta_{2M}, \eta_{2M}). \quad (3.11.2)$$

Also, let θ_C denote the vector such that the four elements will provide for the maximum likelihood value with the parameters constrained such that the sample 10-percent lives are equal,

$$\theta_C = (\beta_{1C}, \eta_{1C}, \beta_{2C}, \eta_{2C}, B10_1 = B10_2). \quad (3.11.3)$$

The B10 notation denotes the 10-percent quantile. Furthermore, the subscripts “C” in this equation are used to make clear that the parameter values differ from the maximum likelihood solutions and instead represent the values that maximize the likelihood of the data consistent with the constraint as listed. The vector θ_M has four degrees-of-freedom while the vector θ_C has three degrees-of-freedom. The null hypothesis may be rejected at the 100α percent confidence level if

$$-2 \times \log \left\{ \max_{B10} \frac{L(\theta_C)}{L(\theta_M)} \right\} > \chi^2_{(1-\alpha;1)}. \quad (3.11.4)$$

Here it is understood that the value chosen for the constraint, B10, is the value that maximizes the value of $L(\theta_C)$. By application of the method, one can either reject or fail to reject the null hypothesis on the basis of the experimental evidence as quantified by the confidence number.

The newly proposed method will now be illustrated using an example. The data to be analyzed are the gear fatigue life data presented in chapter 4 (fig. 4.3.2). Here the data can be considered as example data for purposes of illustration. The parameter values that maximizes the likelihood of the data, that is the maximum likelihood solutions, are $\theta_M = (0.9, 103, 1.0, 377)$. Here, the units for the scale parameters (and therefore for any predicted quantile) are millions of stress cycles. The sample 10-percent life estimates are the values 7.5 and 43.0. The next step is to form the null hypothesis that the 10-percent lives are equal. The value of 10-percent life that maximizes the constrained likelihood will be contained within the interval bound by the maximum likelihood estimates, i.e. the interval [7.5 43.0]. The appropriate value for the 10-percent life constraint can be found by trial-and-error. The maximized likelihood ratio,

$L(\theta C)/L(\theta M)$, as a function of the assumed value for the common 10-percent life is provided in figure 3.11.1. The likelihood ratio values calculated are marked by the symbols, and a spline curve was fit through the results. The ordinate is labeled as the maximized likelihood ratio to emphasize the condition that for each presumed value of a common 10-percent life, the two remaining free parameters were selected to maximize the ratio. The appropriate values for the free parameters were found by producing contour plots of the likelihood ratio as a function of the two free parameters. Such a contour plot for the proposed common 10-percent life equal to 10.0 is provided in figure 3.11.2. The maximized likelihood ratio value for each proposed common 10-percent life as plotted by a symbol on figure 3.11.1 was determined using a contour plot similar to that of figure 3.11.2. As the final step of the calculations, the maximum possible likelihood ratio under the constraint of a common 10-percent life is related to a confidence number using the relation of equation (3.11.3). For a proposed common 10-percent life equal to 10.9, the maximized likelihood ratio is 0.255. Making use of equation (3.11.3), the null hypothesis can be rejected to a confidence level of 91 percent.

To further explore the concept of the proposed method for comparing two populations, an approximate sampling distribution was calculated for the difference of sample 10-percent lives. Two Weibull distributions were defined, each having a 10-percent life of 10.9 but differing shape factors as had provided the maximized likelihood ratio. Monte Carlo simulation was then used to simulate the random sampling process. The number of samples selected and censoring scheme used for the experiments were reproduced by the Monte Carlo scheme. The final result of each Monte Carlo simulation set was an observed difference in sample 10-percent lives. In total 20,000 simulation sets were completed, and the results were analyzed. The approximate sampling distribution for the difference in sample 10-percent lives is provided in figure 3.11.3. The observed difference in experimental 10-percent lives was the value $(43.0-7.5) = 35.5$ (chapter 4). Referring to 3.11.3, by simulation we determine that if the null hypothesis were true, then 10-percent life differences as large as 35.5 are infrequent. By sorting the Monte Carlo simulation results, it was found that the value of 35.5 corresponds to the 91st percentile. The observed difference of such a magnitude would occur due to random sampling effects only nine-percent of the time if the proposed null-hypothesis were true. The analysis indicates that the random sampling effect is an unlikely explanation for the experimentally observed differences in 10-percent life. Therefore, the null hypothesis is rejected, and it is declared that the observed difference in the estimated 10-percent lives is a statistically significant difference. Note that if a higher confidence level (say 95 percent confidence) was desired, then we would fail to reject the null hypothesis.

3.12 Conclusions and Recommendations

A practitioner faced with the task of statistical analysis of fatigue and reliability data needs to select the most appropriate method from the many that have been proposed. For gear surface fatigue data, probability density functions are typically skewed right, samples sizes usually range from 10 to 40, and often censoring is limited to Type I censoring. In addition, the estimation of the distribution parameters is often a means to an end, the goal being the estimation of quantiles of the cumulative distribution function. Another end goal is to be able to compare two populations with (potentially) differing fatigue life distributions. The studies and developments of this chapter were done with these ideas and end goals in mind. As a result of the appropriately focused studies four specific issues were resolved, and recommendations were provided. The significant contributions of the work described in this chapter now follows:

1. Software for the calculation of parameter estimates and confidence intervals were developed and validated.
2. The usual practice of describing gear fatigue data using the 2-parameter Weibull distribution rather than the more general 3-parameter Weibull distribution was critically examined. Although the experimental evidence and theoretical considerations both suggest that the true value of the threshold parameter is not identically zero, the presumption of a zero-valued threshold is a good

approximation. It is recommended to model gear surface fatigue life distributions with a 2-parameter Weibull distribution and, thereby, as having a zero-valued threshold parameter.

3. Three methods for determining distribution parameters from sample data (two regression-based methods and a maximum likelihood based method) were evaluated for accuracy and precision. It is recommended to estimate the parameters of the Weibull distribution to describe gear surface fatigue life by using the maximum likelihood method.
4. A new method is proposed and developed for comparing two datasets. The purpose of the new method is to detect the existence of statistically significant differences in fatigue life properties. The method compares two datasets based on a selected quantile of the fatigue life cumulative distribution functions. A null hypothesis is set forth that a chosen quantile of the two populations are equal. By application of the method, one can either reject or fail-to-reject the null hypothesis on the basis of the experimental evidence as quantified by a confidence number.

In conducting the studies presented in this chapter, some ideas have come to mind that could extend the methods provided here. These ideas were beyond the scope of the present effort, but they are suggested as ideas that might provided valuable results for future experimental work to evaluate gear surface fatigue lives. A list of suggested topics for future research is now provided.

- a. In the present work, the statistical tools were evaluated and developed with the end goal in mind the statistical description and inference of gear surface fatigue data. Modern statistics also offer powerful tools for test planning to help optimize testing variables such as number of samples required and censoring schemes. Results from appropriately focused studies for gear surface fatigue research would likely yield valuable results.
- b. In this work, a critical examination of the usual presumption of a zero-valued threshold parameter was completed. Two options were presented as alternate choices: either the threshold is presumed to be identically zero or it must be estimated, separately, for each dataset. Of these two choices, the zero-valued presumption was selected as the preferred choice. However, one could propose still another choice by making some other presumption about the value of the threshold parameter. The presumption should be guided by historical data. For example, one could presume that the threshold parameter is some proportion of the scale parameter, and the ratio of the threshold to scale could be modeled as a constant value for all distributions. The value for the ratio could be determined by analysis of a representative set of gear surface fatigue experiments. For example, consider if one were to make use of ten experiments. Fitting these data using the zero-valued presumption and the Weibull distribution requires 20 free parameters. Discarding the zero-valued presumption and instead estimating a different threshold for each experiment requires 30 free parameters. The alternative choice, proposed here, would be to presume that the thresholds are non-zero and that, from experiment to experiment, the ratio of threshold to scale is a constant. Such an approach would require estimating 21 free parameters. The method of Tallian (ref. 3.35) would be another alternative for determining a historical-based normalized threshold. The very recent work of Shimizu may also offer additional insight (ref. 3.40).
- c. In this work, a new method is proposed for comparing two datasets. The method is a way to evaluate the null hypothesis that the two datasets have fatigue lives that are equal at some chosen quantile of interest. This method could be extended to test other hypothesis, for example that the quantiles of interest are separated by some specified value. In fact, the method could be extended to create a profile likelihood curve for the sample statistic “difference in X-percent lives,” where “X” is any quantile of interest. Determining this curve would require a large number of calculations. Recall that each point on the profile likelihood curve represents the peak of another curve like the one illustrated in figure 3.11.1. Furthermore, each point on the curve of figure 3.11.1 is itself a peak value obtained from a contour plot like figure 3.11.2. Still, the profile likelihood curve can be determined, and the approach would be feasible if the appropriate interactive or automated computing tools were to be developed.

References

- 3.1. Abernathy, R.; Breneman, J.; Medlin, C.; Reinman, G., *Weibull Analysis Handbook*, Air Force Wright Aeronautical Laboratories Report AFWAL-TR-83-2079, (1983). Available from National Technical Information Service, Washington, DC.
- 3.2. Hallinan, A., "A Review of the Weibull Distribution," *J. of Quality Technology*, **25** [2], (1993).
- 3.3. Nelson, W., *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, John Wiley and Sons, New York, (1990).
- 3.4. Jacquelin, J., "A Reliable Algorithm for the Exact Median Rank Function," *IEEE Trans. On Electrical Insulation*, **28** [2], (1993).
- 3.5. Harter, H.; Moore, A., "Maximum-Likelihood Estimation of the Parameters of Gamma and Weibull Populations from Complete and from Censored Samples," *Technometrics*, **7** [4], (1965).
- 3.6. Soman, K.; Misra, K., "A Least Square Estimation of Three Parameters of a Weibull Distribution," *Microelectronics Reliability*, **32** [3], (1992).
- 3.7. Hirose, H., "Maximum Likelihood Estimation in the 3-parameter Weibull Distribution," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996) and ERRATUM, **3** [2], (1996).
- 3.8. Keats, J.; Lawrence, F.; Wang, F., "Weibull Maximum Likelihood Parameter Estimates with Censored Data," *J. of Quality Technology*, **29** [1], (1997).
- 3.9. Krantz, T., "Software for Statistical Analysis of Weibull Distributions With Application to Gear Fatigue Data - User's Manual with Verification," *NASA TM-2002-211109, ARL-TR-1338*, (2002).
- 3.10. Bergman, B., "Estimation of Weibull Parameters Using a Weight Function," *J. of Materials Science Letters*, **5** [6], (1986).
- 3.11. Pieracci, A., "Parameter Estimation for Weibull Probability Distribution Function of Initial Fatigue Quality," *AIAA Journal*, **33** [9], (1995).
- 3.12. Ross, R., "Bias and Standard Deviation due to Weibull Parameter Estimation for Small Data Sets," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996).
- 3.13. Cacciari, M.; Mazzanti, G.; Montanari, G., "Comparison of Maximum-Likelihood Unbiasing Methods for the Estimation of the Weibull Parameters," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996).
- 3.14. Jacquelin, J., "Inference of Sampling on Weibull Parameter Estimation," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [6], (1996).
- 3.15. Montanari, G. et al., "In Search of Convenient Techniques for Reducing Bias in the Estimation of Weibull Parameters for Uncensored Tests," *IEEE Trans. On Dielectrics and Electrical Insulation*, **4** [3], (1997).
- 3.16. Hirose, H., "Bias Correction for the Maximum Likelihood Estimates in the Two-parameter Weibull Distribution," *IEEE Trans. On Dielectrics and Electrical Insulation*, **6** [1], (1999).
- 3.17. Escobar, L.; Meeker, W., "Statistical Prediction Based on Censored Life Data," *Technometrics*, **41** [2], (1999).
- 3.18. Jeng, S.; Meeker, W., "Comparisons of Approximate Confidence Interval Procedures for Type I Censored Data," *Technometrics*, **42** [2], (2000).
- 3.19. Coddington, P., "Using Random Number Generators," (1995). (Available online <http://www.npac.syr.edu/projects/random/index.html>) Accessed Aug. 6, 2001.
- 3.20. Fishman, G.; Moore, L., "A Statistical Evaluation of Multiplicative Random Number Generators With Modulus 2311," *J. of the American Statistical Society*, **77**, (1982).
- 3.21. Fishman, G.; Moore, L., "An Exhaustive Analysis of Multiplicative Congruential Random Number Generators with Modulus 231-1," *SIAM J. on Scientific and Statistical Computing*, **7**, (1986).
- 3.22. Fushimi, M., "Random Number Generation with the Recursion $X_t = X_{t-3p} \oplus X_{t-3q}$," *J. of Computational and Applied Mathematics*, **31**, (1990).
- 3.23. Cohen, A., "The Reflected Weibull Distribution," *Technometrics*, **15** [4], (1973).

- 3.24. Engelhardt, M.; Bain, L.. "Simplified Statistical Procedures for the Weibull or Extreme-value Distribution," *Technometrics*, **19** [3], (1977).
- 3.25. Cleveland, W., *Visualizing Data*. Hobart Press. New Jersey, (1993).
- 3.26. Coy, J.; Townsend, D.; Zaretsky, E., "Analysis of Dynamic Capacity of Low-Contact-Ratio Spur Gears Using Lundberg-Palmgren Theory," *NASA TN-D-8029*, (1975).
- 3.27. Townsend, D.; Bambergher, E.; Zaretsky, E., "Comparison of Pitting Fatigue life of Ausforged and Standard Forged AISI M-50 and AISI 9310 Spur Gears," *NASA TN-D-8030*, (1975).
- 3.28. Townsend, D.; Coy, J.; Zaretsky, E., "Experimental and Analytical Load-Life Relation for AISI 9310 Steel Spur Gears," *NASA TM-X-73590*, (1977).
- 3.29. Townsend, D.; Zaretsky, E., "Effect of Shot Peening on Surface Fatigue Life of Carburized and Hardened AISI 9310 Spur Gears," *NASA TP-2047*, (1982).
- 3.30. Scibbe, W.; Townsend, D.; Aron, P., "Effect of Lubricant Extreme-Pressure Additives on Surface Fatigue Life of AISI 9310 Spur Gears," *NASA TP-2408*, (1984).
- 3.31. Townsend, D.; Zaretsky, E., "Effect of Five Lubricants on Life of AISI 9310 Spur Gears," *NASA TP-2419*, (1985).
- 3.32. Townsend, D., "Surface Fatigue Life and Failure Characteristics of EX-53, CBS 1000M, and AISI 9310 Gear Materials," *NASA TP-2513*, (1985).
- 3.33. Townsend, D., "Surface Fatigue Life of Noninvolute, Low-Contact Ratio Gears," *NASA TM-103116*, (1990).
- 3.34. Townsend, D.; Shimski, J., "Evaluation of the EHL Film Thickness and Extreme Pressure Additives on Gear Surface Fatigue Life," *NASA TM-106663*, (1994).
- 3.35. Tallian, T., "Weibull Distribution of Rolling Contact Fatigue Life and Deviations There From," *ASLE Transactions*, **5**, (1962).
- 3.36. Doganaksoy, N.; Schmee, J., "Comparisons of Approximate Confidence Intervals for Distributions Used in Life-Data Analysis," *Technometrics*, **35** [2], (1993).
- 3.37. Meeker, W.; Escobar, L., *Statistical Methods for Reliability Data*, John Wiley and Sons, New York (1998).
- 3.38. Johnson, L., *The Statistical Treatment of Fatigue Experiments*, Elsevier, New York, (1964).
- 3.39. McCool, J., "Evaluating Weibull Endurance Data by the Method of Maximum Likelihood," *ASLE Transactions*, **13** [3], (1970).
- 3.40. Shimizu, S., "Fatigue Limit Concept and Life Prediction Model for Rolling Contact Machine Elements," *Tribology Transactions*, **45** [1], (2002).

TABLE 3.1.1.—VALIDATION OF MONTE CARLO SCHEME

(a) 10 random samples per simulation set

Sampled distribution	Method for calculating sampling distribution	Method for fitting distribution parameters	10th percentile of the sampling distribution	Median of sampling distribution	90th percentile of the sampling distribution
Normal	Exact theory	N/A	0.736	0.897	1.058
Normal	Monte-Carlo	sample mean formula	0.738	0.898	1.058
Weibull	Monte-Carlo	2-parameter least-squares regression	0.730	0.894	1.059
Weibull	Monte-Carlo	3-parameter least-squares regression	0.705	0.887	1.060
Weibull	Monte-Carlo	maximum likelihood	0.736	0.899	1.063

(b) 30 random samples per simulation set

Sampled distribution	Method for calculating sampling distribution	Method for fitting distribution parameters	10th percentile of the sampling distribution	Median of sampling distribution	90th percentile of the sampling distribution
Normal	Exact theory	N/A	0.804	0.897	0.990
Normal	Monte-Carlo	sample mean formula	0.805	0.897	0.990
Weibull	Monte-Carlo	2-parameter least-squares regression	0.800	0.895	0.990
Weibull	Monte-Carlo	3-parameter least-squares regression	0.791	0.896	0.997
Weibull	Monte-Carlo	maximum likelihood	0.805	0.897	0.990

TABLE 3.3.2.—MAXIMUM AND MINIMUM VALUES OBTAINED FOR CHOSEN PERCENTILES OF APPROXIMATE SAMPLING DISTRIBUTIONS AS OBTAINED FROM 60 MONTE CARLO STUDIES*

(a) 2-parameter least-squares regression

Statistic of interest	Random sampling method	5th percentile (max., min.)	50th percentile (max., min.)	95th percentile (max., min.)
10-percent life	MCM	0.0288, 0.0313	0.1296, 0.1332	0.3518, 0.3656
10-percent life	GFSR	0.0289, 0.0311	0.1298, 0.1334	0.3507, 0.3634
50-percent life	MCM	0.4113, 0.4208	0.7263, 0.7353	1.1510, 1.1758
50-percent life	GFSR	0.4110, 0.4224	0.7269, 0.7358	1.1561, 1.1743

(b) 3-parameter least-squares fits

Statistic of interest	Random sampling method	5th percentile (max., min.)	50th percentile (max., min.)	95th percentile (max., min.)
10-percent life	MCM	0.0341, 0.0371	0.1454, 0.1495	0.3965, 0.4124
10-percent life	GFSR	0.0346, 0.0370	0.1463, 0.1505	0.3982, 0.4092
50-percent life	MCM	0.3792, 0.3896	0.7123, 0.7230	1.2011, 1.1734
50-percent life	GFSR	0.3786, 0.3890	0.7137, 0.7218	1.1782, 1.2024

MCM - multiplicative congruential method for random samples

GFSR- generalized feedback shift register method for random samples

- Note: each Monte Carlo study comprised of 20,000 simulation sets

TABLE 3.6.1.—DEFINING PARAMETERS OF FOUR WEIBULL DISTRIBUTIONS HAVING SIMILAR LIVES
WITHIN THE RANGE OF [10 TO 90] CUMULATIVE PERCENT FAILED (ALSO SEE FIG. 3.6.1)

Distribution	Shape	Scale	Threshold	10-percent life
1	2.00	1.00	0.0	0.325
2	1.70	0.90	0.1	0.340
3	1.53	0.79	0.2	0.381
4	1.25	0.67	0.3	0.411

TABLE 3.8.1.—RESULTS OF MONTE CARLO SIMULATION OF LIFE
TESTING OF SAMPLES FROM A WEIBULL POPULATION, WHILE
EMPLOYING A CENSORING TIME EQUAL TO 1.5 TIMES THE
TRUE 10-PERCENT LIFE, SHOWING NUMBER
OF FAILURES FROM 30 TOTAL SAMPLES

Number of failures	Number of occurrences*
2	614
3	1228
4	2290
5	3098
6	3512
7	3361
8	2549
9	1672
10	921
11	470
12	187
13	66
14 or more	32

* Note: Total number of simulation sets was 20,000.

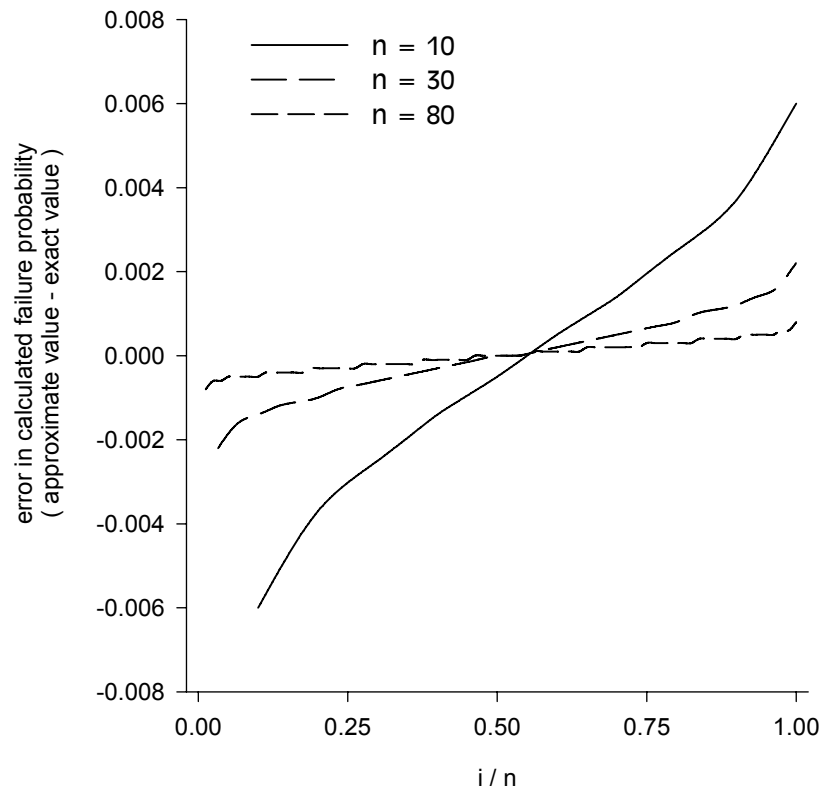


Figure 3.2.1.—Difference between approximate calculations (eq. (3.2.2)) and exact calculations for the median rank cumulative failure probability for ranked item ‘i’ from a dataset of size ‘n’.

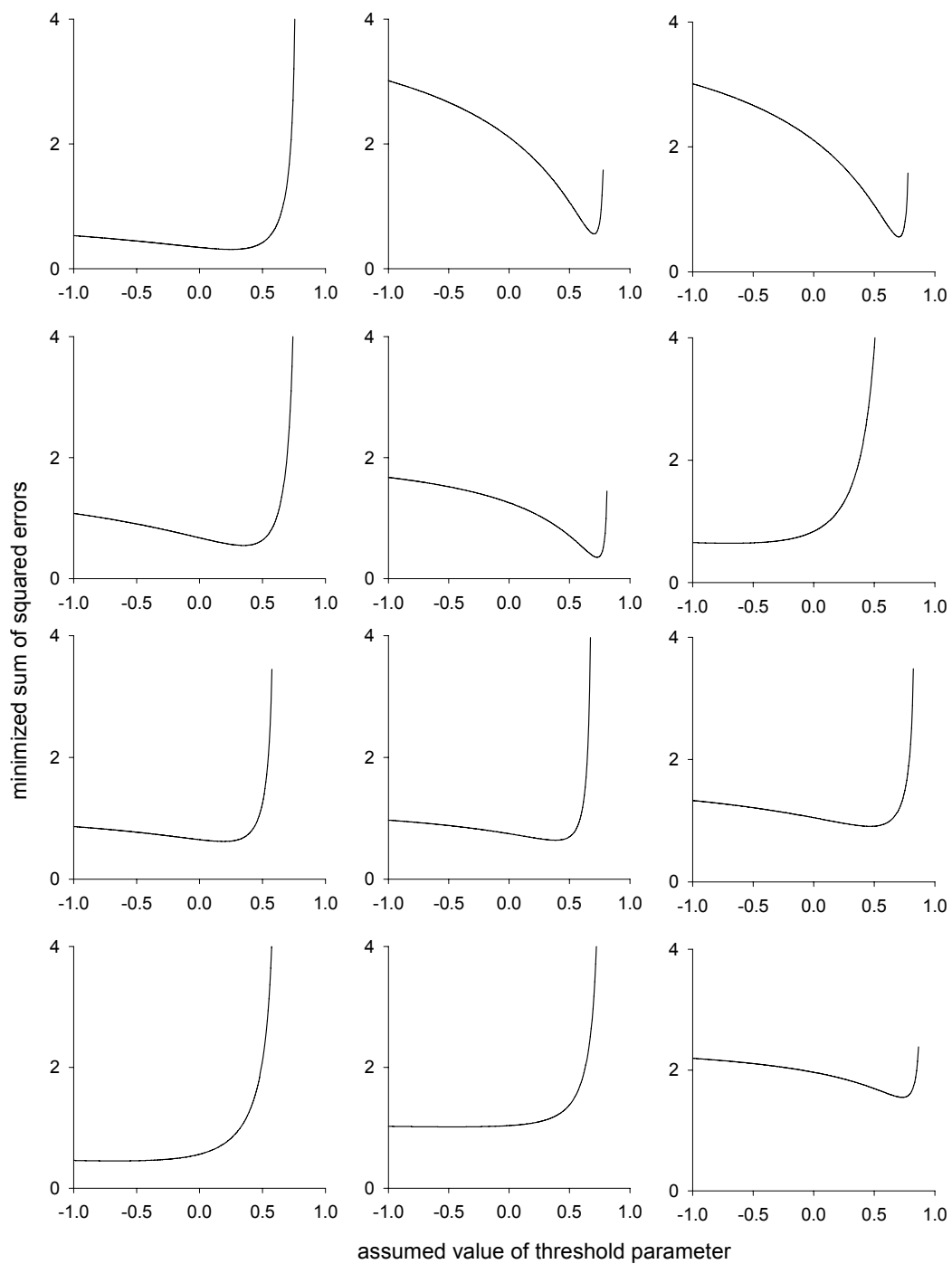


Figure 3.2.2.—Sum of squared errors as a function of assumed values for the threshold parameter for twelve Monte Carlo simulations using least-squares regression to fit the scale and shape parameters.

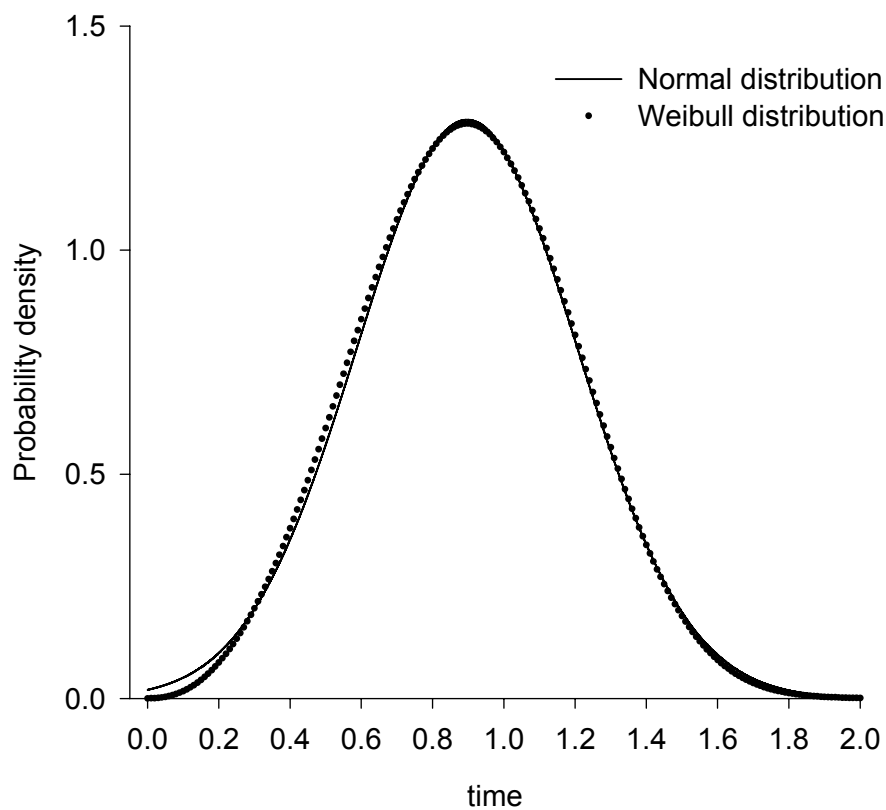


Figure 3.3.1.—Normal and Weibull probability distribution functions used to validate Monte-Carlo simulation scheme and codes. The Weibull distribution parameters were selected to closely approximate a normal distribution.

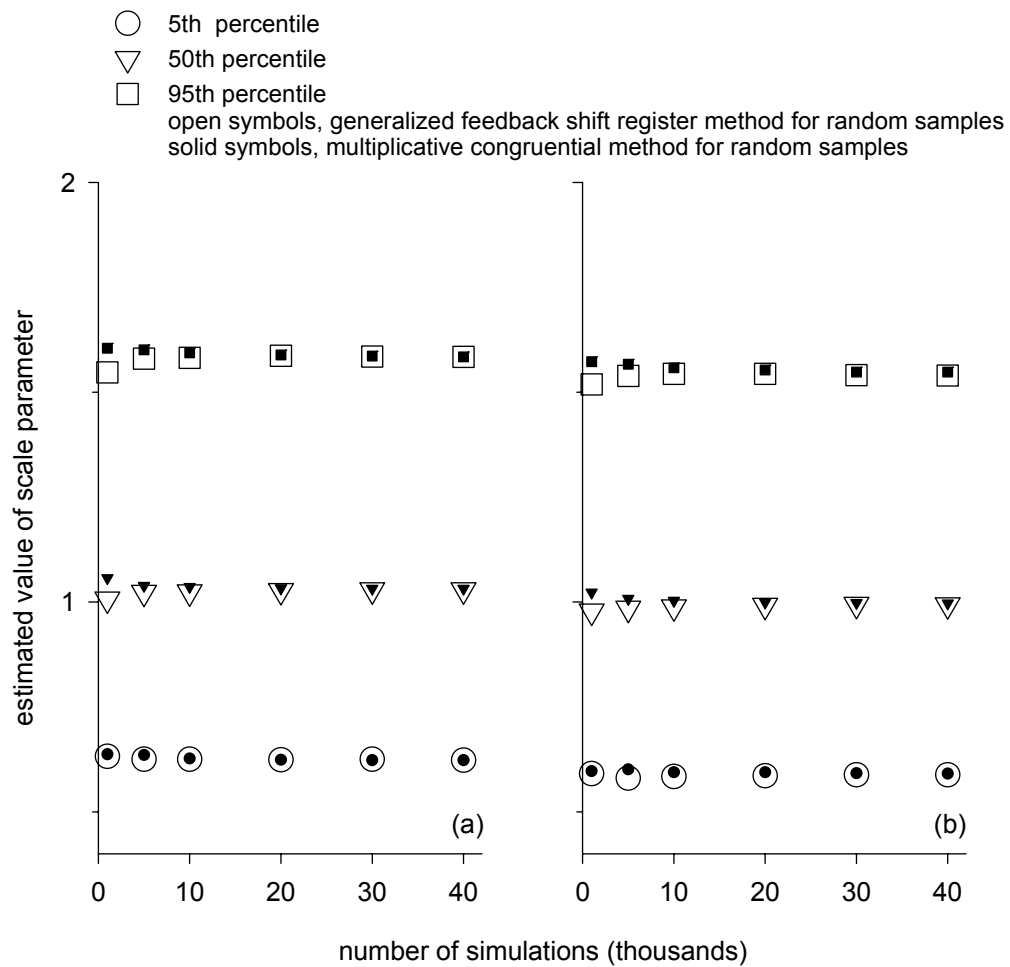


Figure 3.3.2.—Sampling distribution values for scale parameter estimates as a function of the number of Monte Carlo simulation sets. Each simulation set made use of 10 randomly drawn samples from a Weibull distribution with true parameters of scale = 1.0, shape = 1.2, and threshold = 0.0. (a) 2-parameter least-squares regression. (b) 3-parameter least-squares regression.

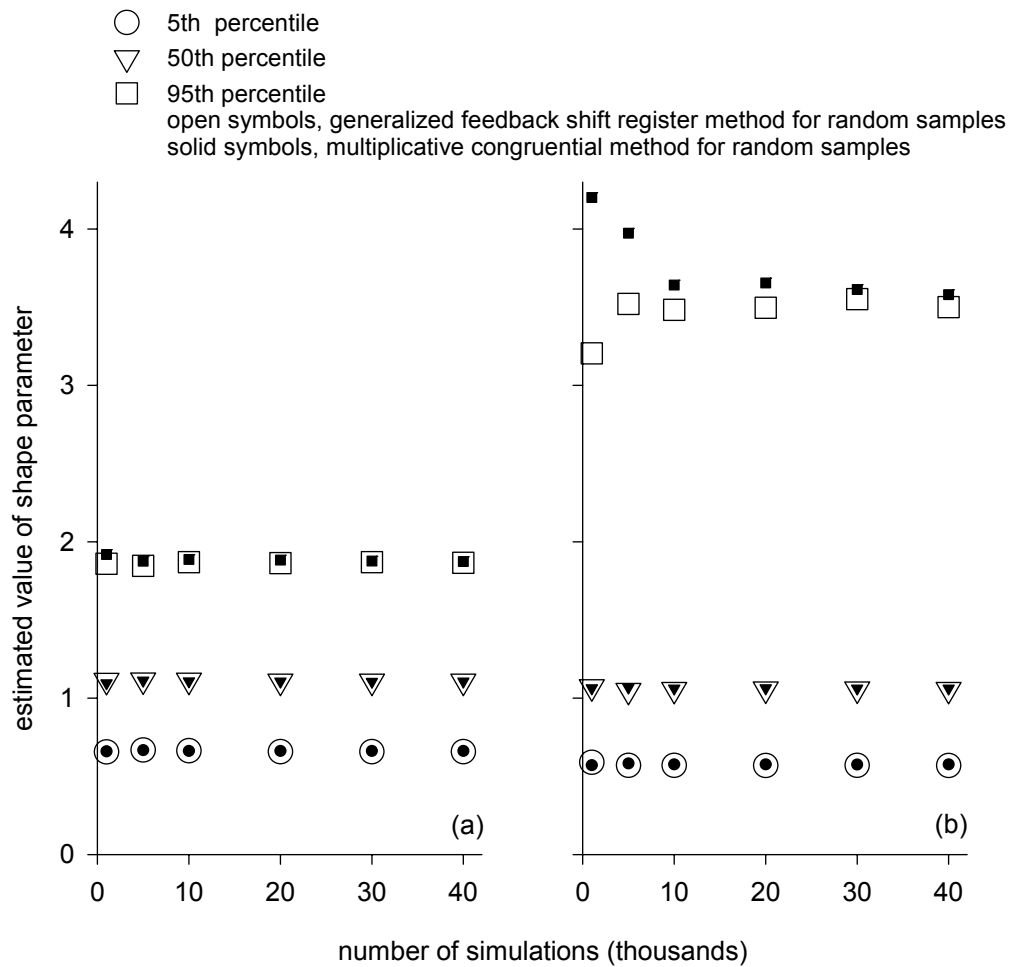


Figure 3.3.3.—Sampling distribution values for shape parameter estimates as a function of the number of Monte Carlo simulation sets. Each simulation set made use of 10 randomly drawn samples from a Weibull distribution with true parameters of scale = 1.0, shape = 1.2, and threshold = 0.0. (a) 2-parameter least-squares regression. (b) 3-parameter least-squares regression.

- 5th percentile
- ▽ 50th percentile
- 95th percentile
- open symbols, generalized feedback shift register method for random samples
- solid symbols, multiplicative congruential method for random samples

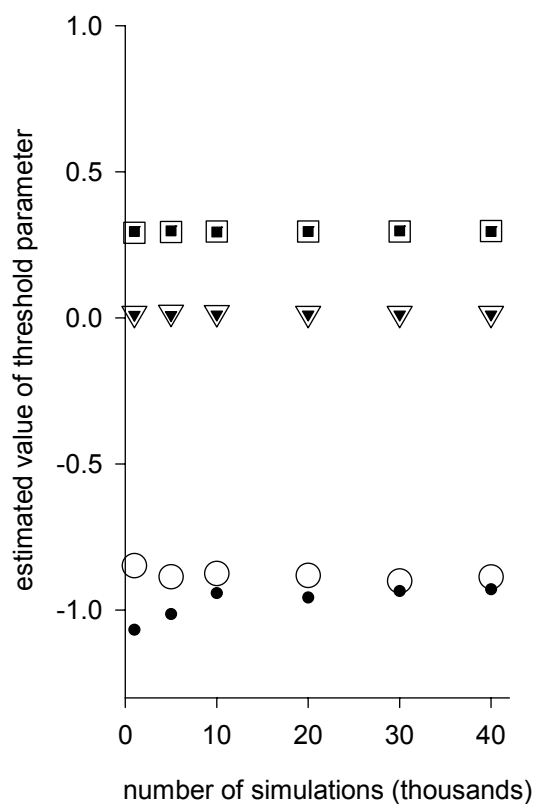


Figure 3.3.4.—Sampling distribution values for threshold parameter estimates as a function of the number of Monte Carlo simulation sets using 3-parameter least-squares regression. Each simulation set made use of 10 randomly drawn samples from a Weibull distribution with true parameters of scale = 1.0, shape = 1.2, and threshold = 0.0.

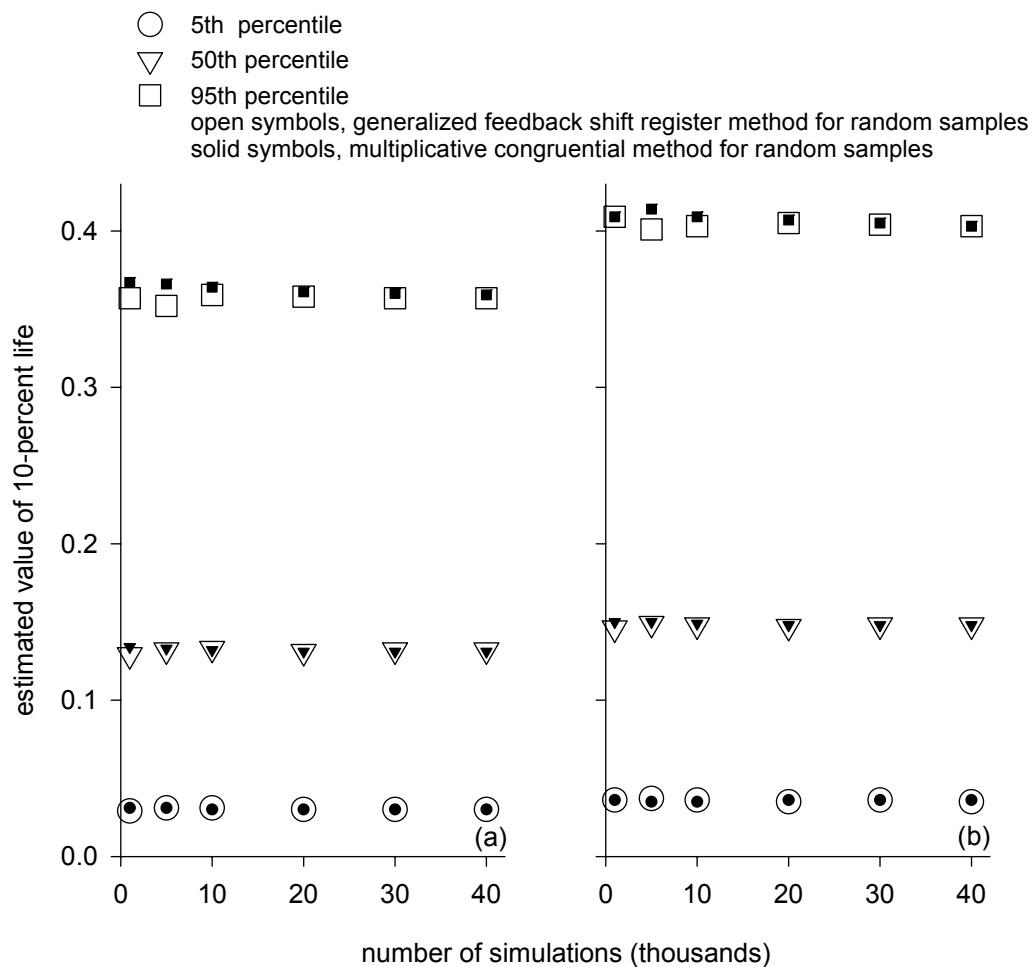


Figure 3.3.5—Sampling distribution values for 10-percent life estimates as a function of the number of Monte Carlo simulation sets. Each simulation set made use of 10 randomly drawn samples from a Weibull distribution with true parameters of scale = 1.0, shape = 1.2, and threshold = 0.0. (a) 2-parameter least-squares regression. (b) 3-parameter least-squares regression.

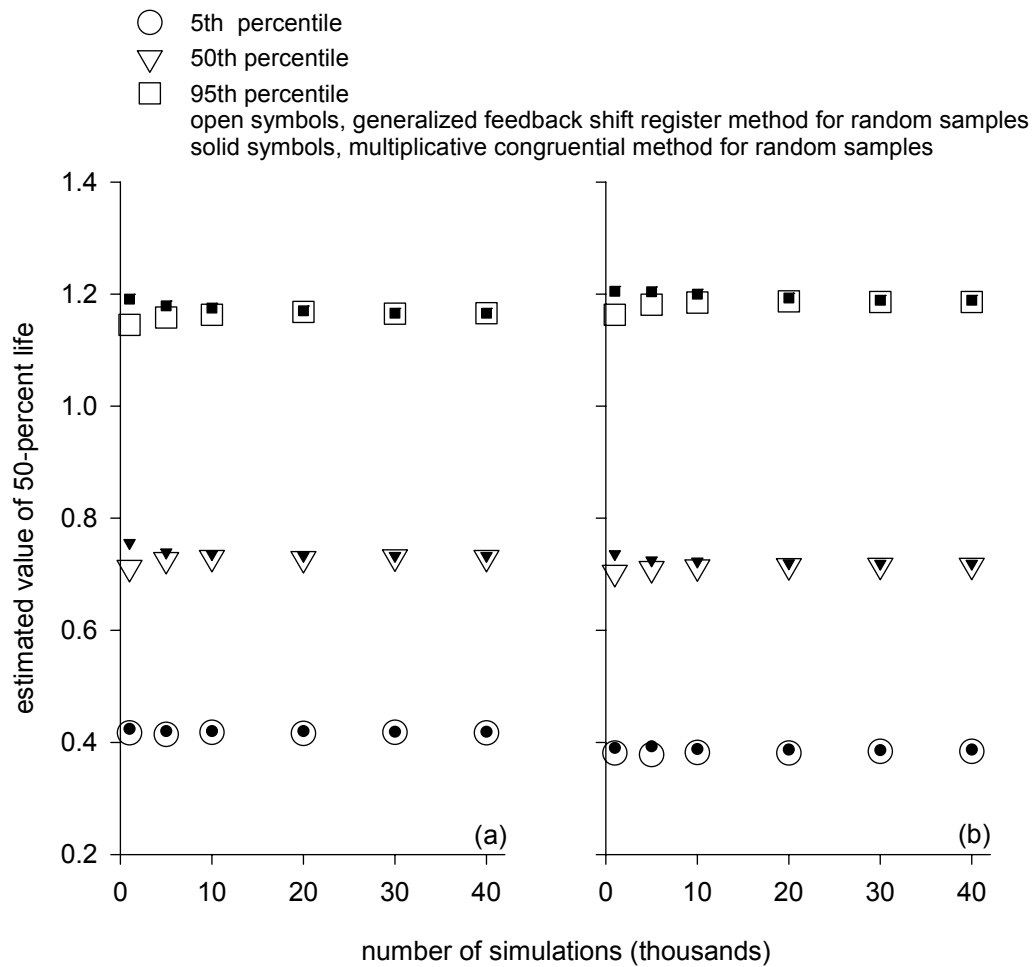


Figure 3.3.6.—Sampling distribution values for 50-percent life estimates as a function of the number of Monte Carlo simulation sets. Each simulation set made use of 10 randomly drawn samples from a Weibull distribution with true parameters of scale = 1.0, shape = 1.2, and threshold = 0.0. (a) 2-parameter least-squares regression. (b) 3-parameter least-squares regression.

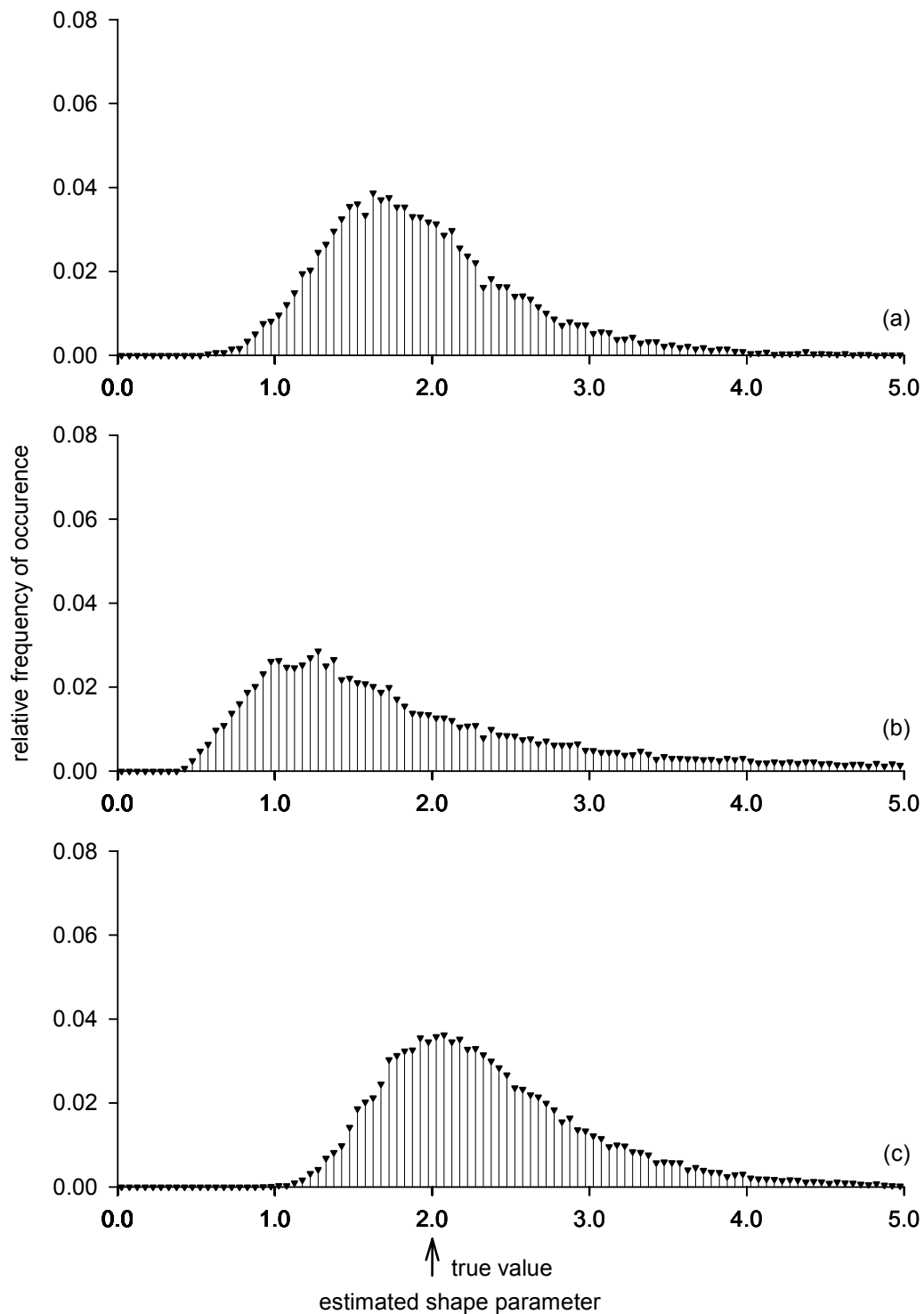


Figure 3.4.1—Approximate sampling distributions for the shape parameter for the case of 10 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method. (c) Maximum likelihood method.

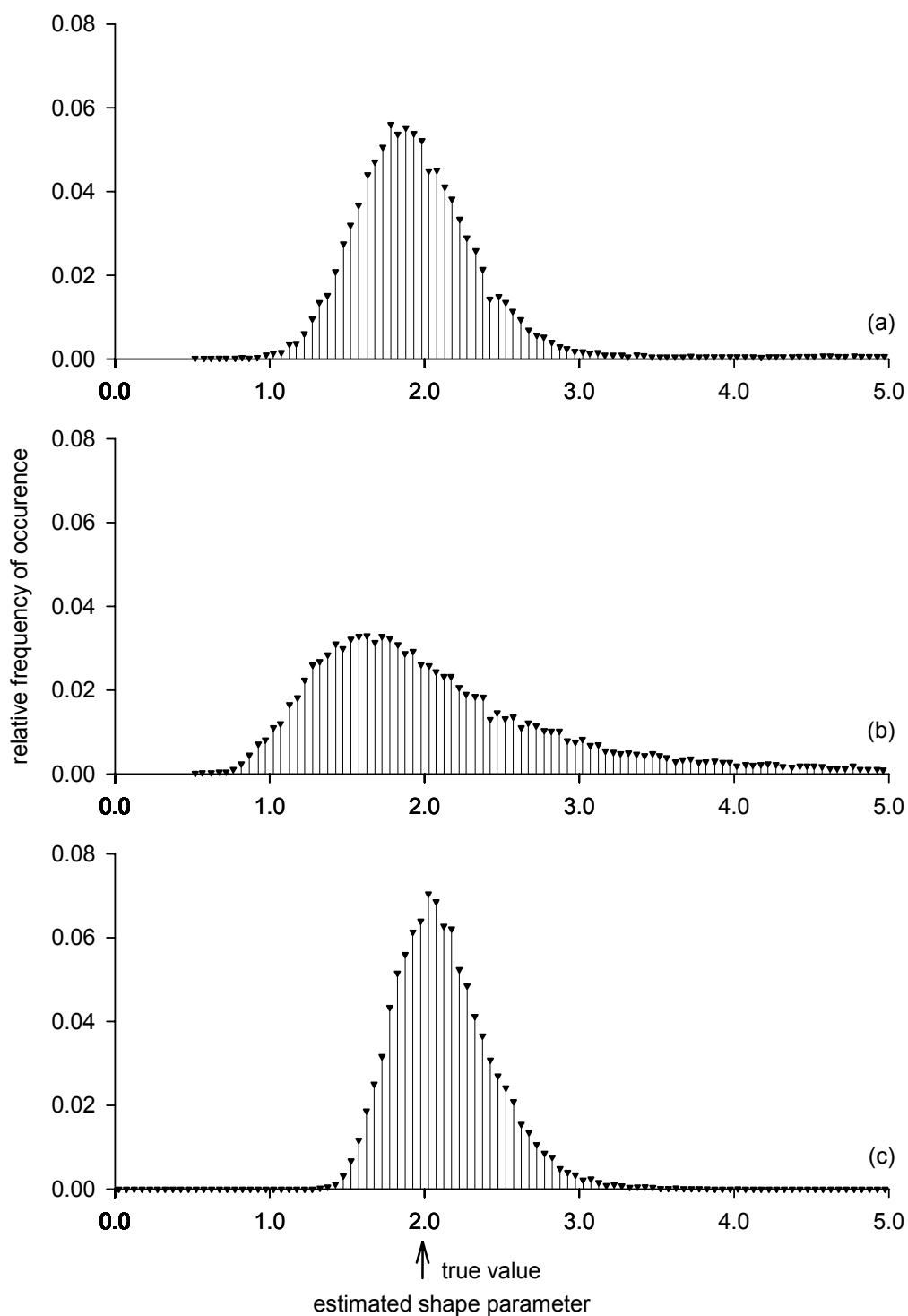


Figure 3.4.2.—Approximate sampling distributions for the shape parameter for the case of 30 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method. (c) Maximum likelihood method.

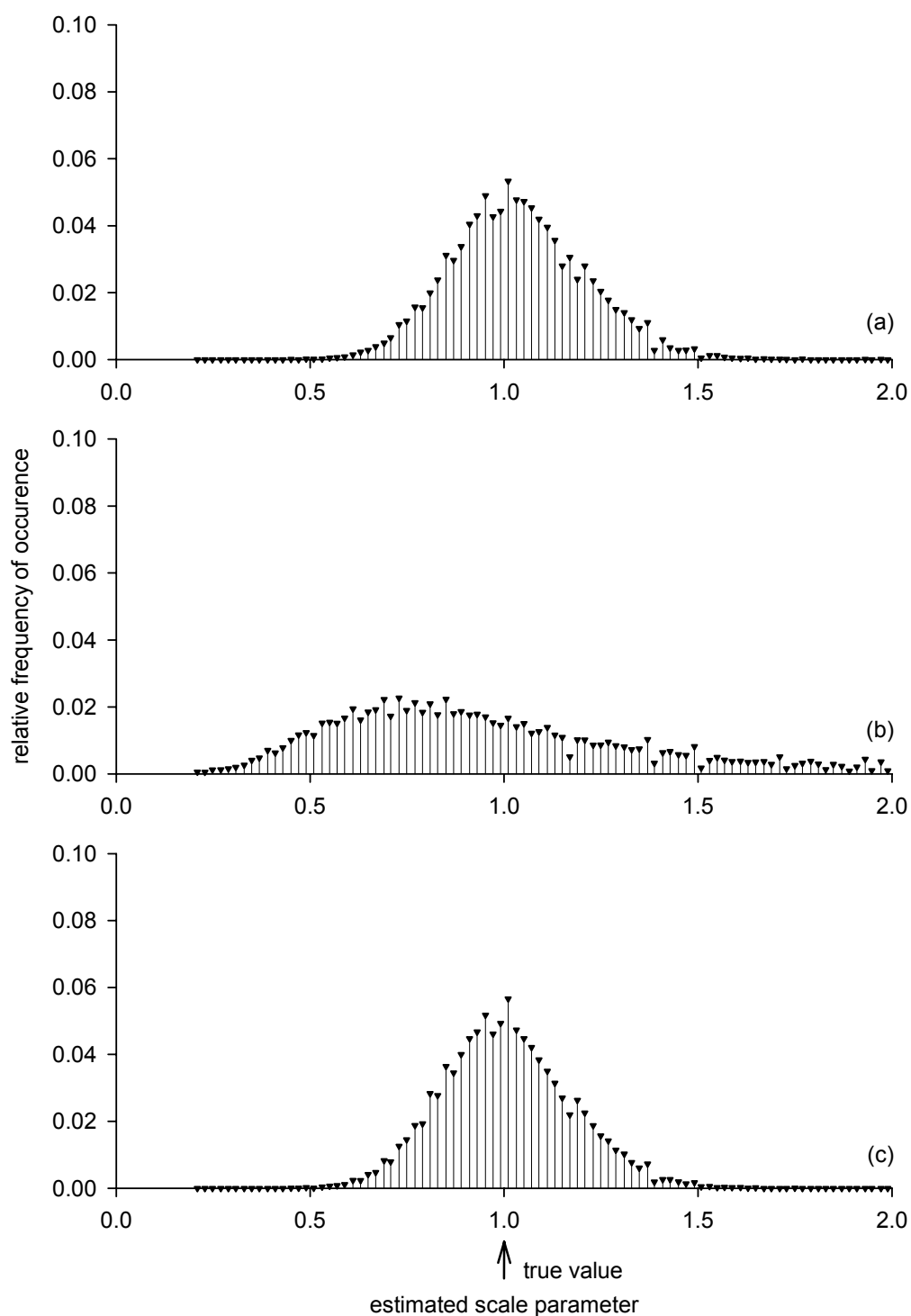


Figure 3.4.3.—Approximate sampling distributions for the scale parameter for the case of 10 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method. (c) Maximum likelihood method.

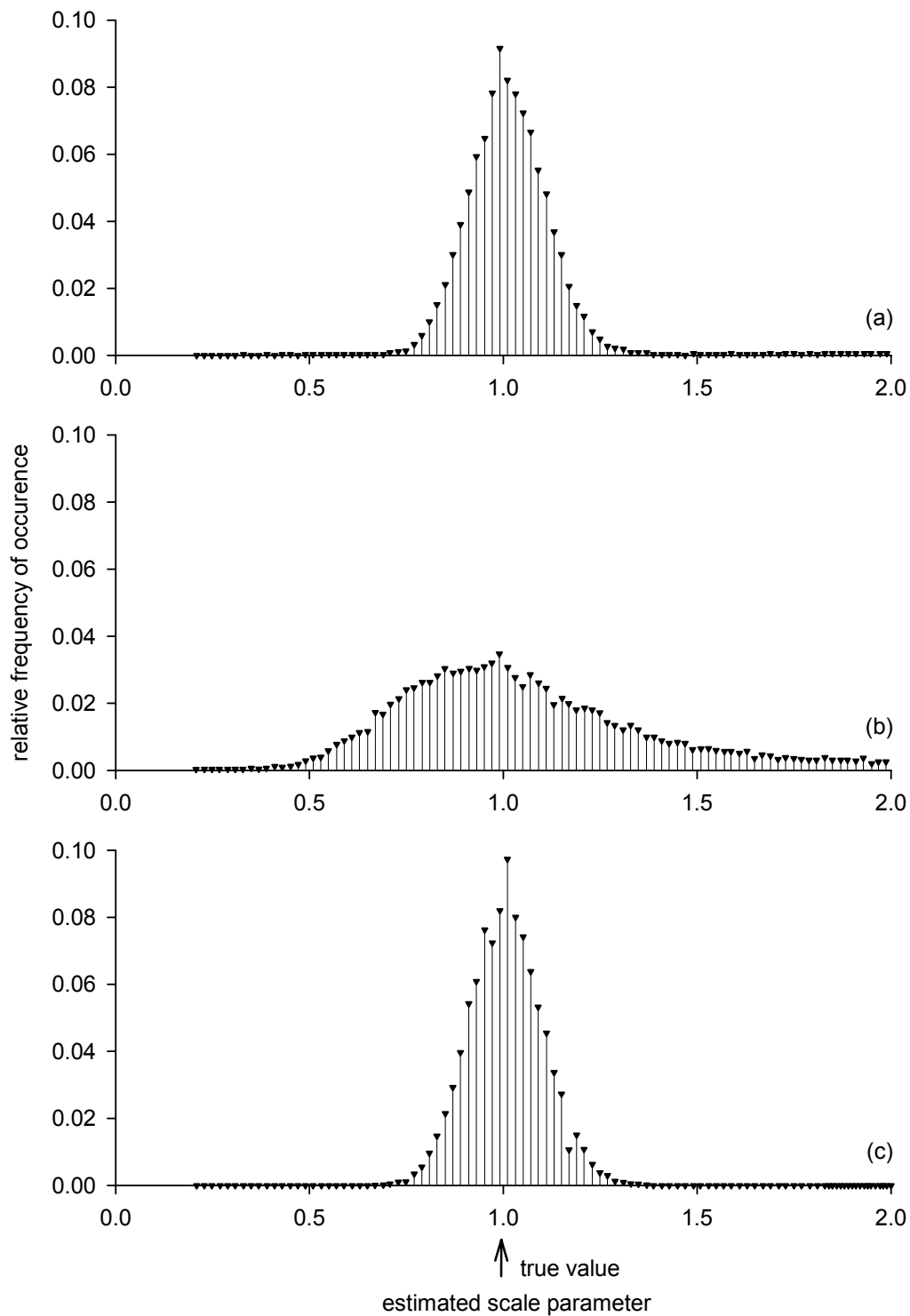


Figure 3.4.4.—Approximate sampling distributions for the scale parameter for the case of 30 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method. (c) Maximum likelihood method.

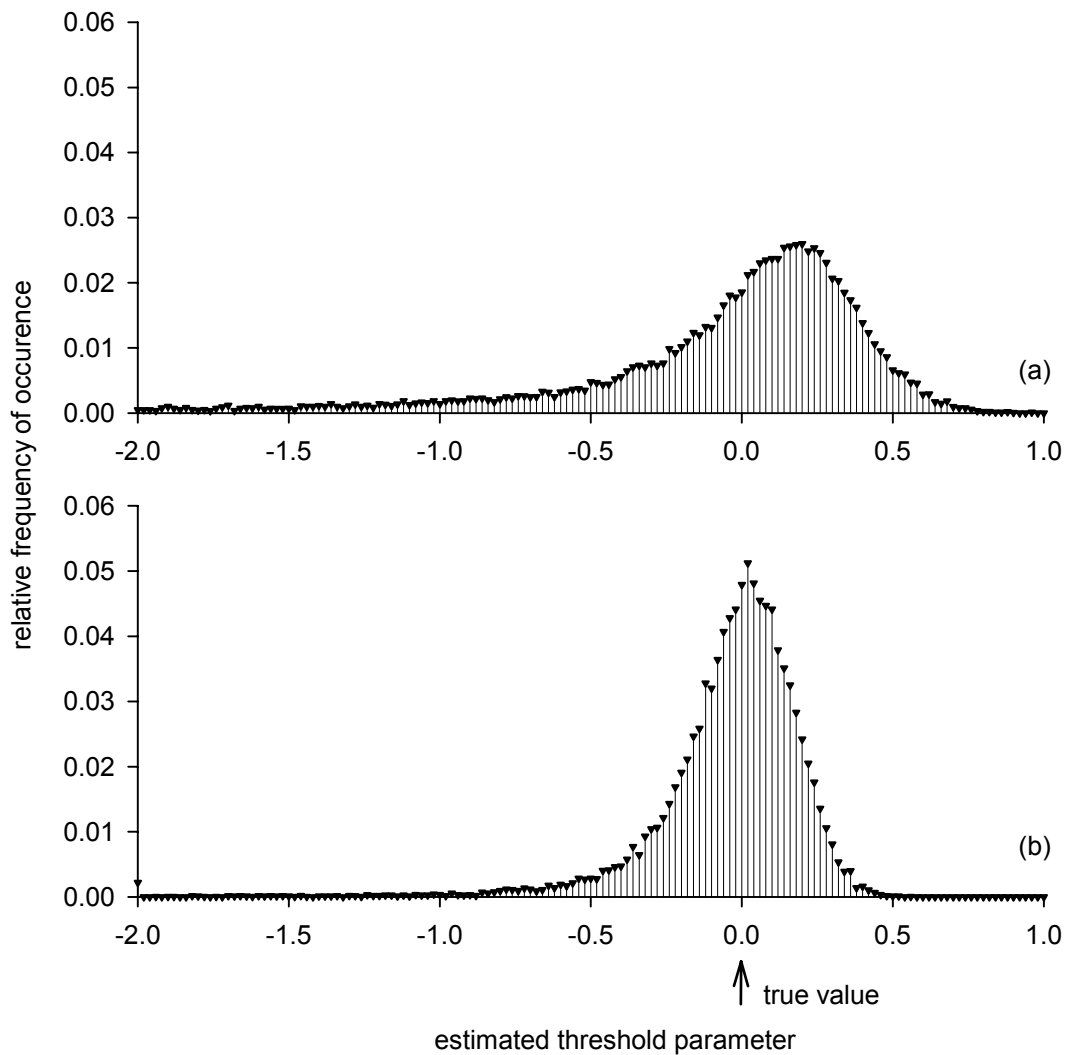


Figure 3.4.5.—Approximate sampling distributions for the threshold parameter for the 3-parameter regression method using random samples selected from a Weibull distribution. (a) Case of 10 random samples per estimate. (b) Case of 30 random samples per estimate.

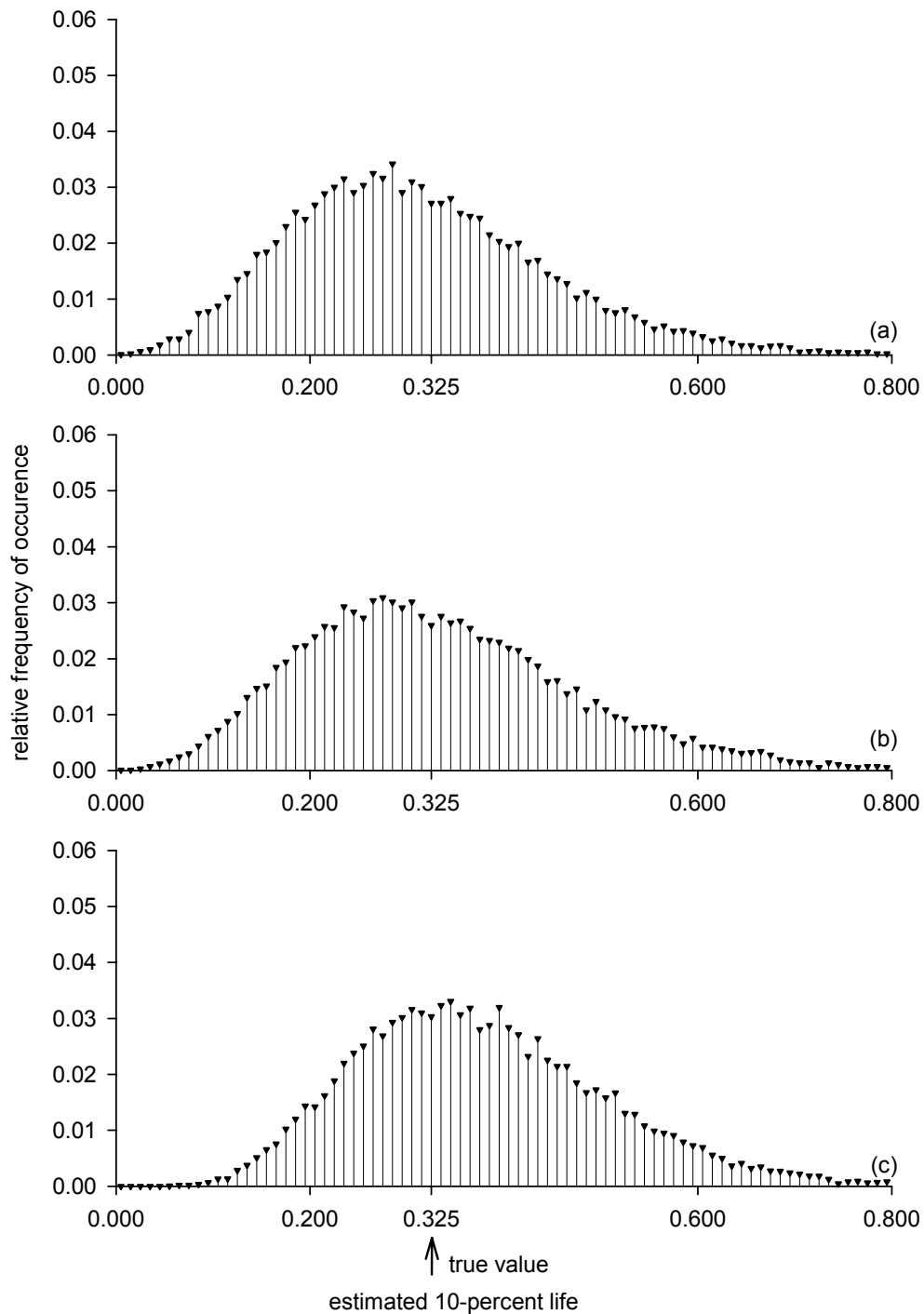


Figure 3.4.6.—Approximate sampling distributions for 10-percent life estimates for the case of 10 random samples selected from a Weibull distribution.
 (a) 2-parameter regression method. (b) 3-parameter regression method.
 (c) Maximum likelihood method.

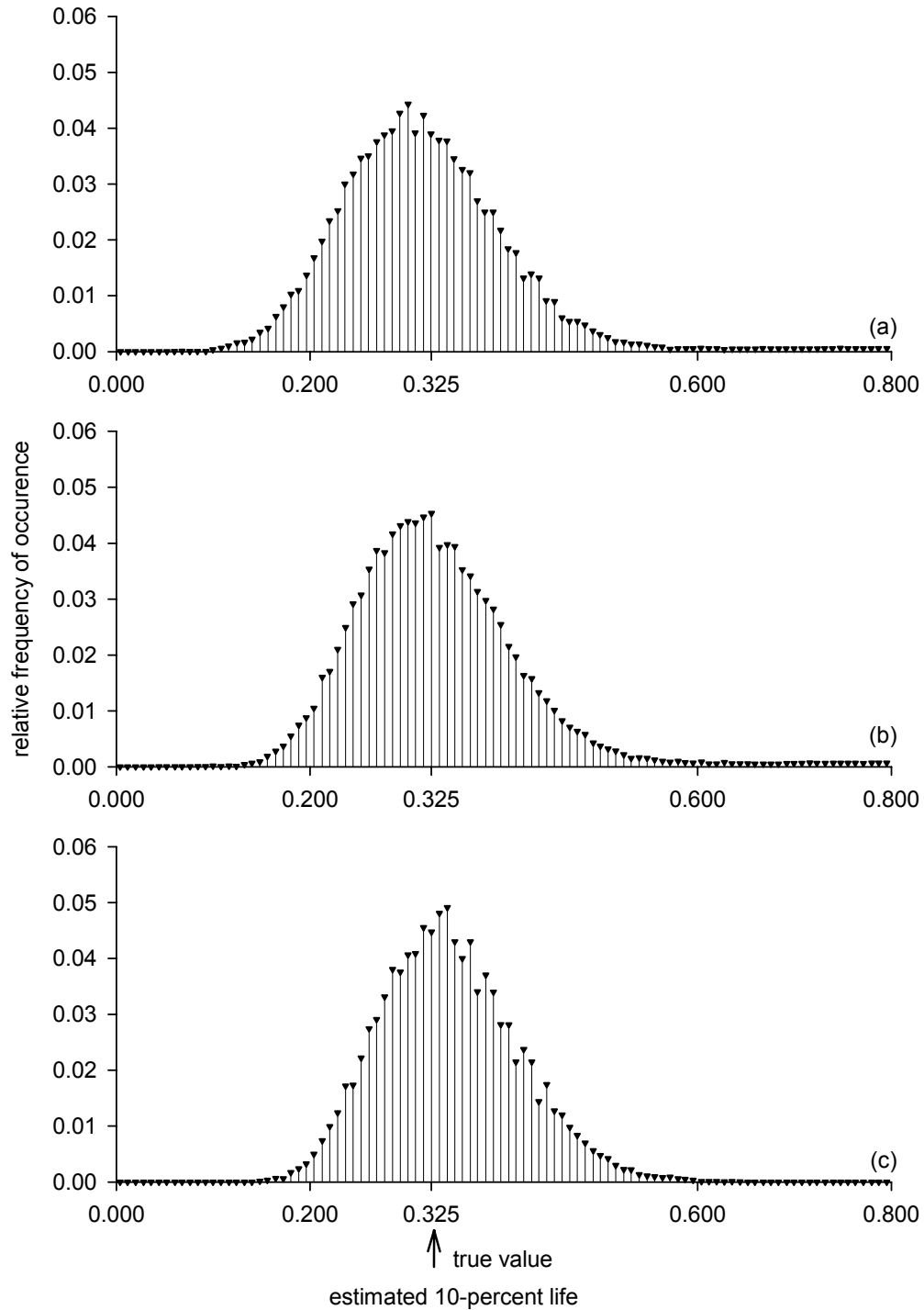


Figure 3.4.7.—Approximate sampling distributions for 10-percent life estimates for the case of 30 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method. (c) Maximum likelihood method.

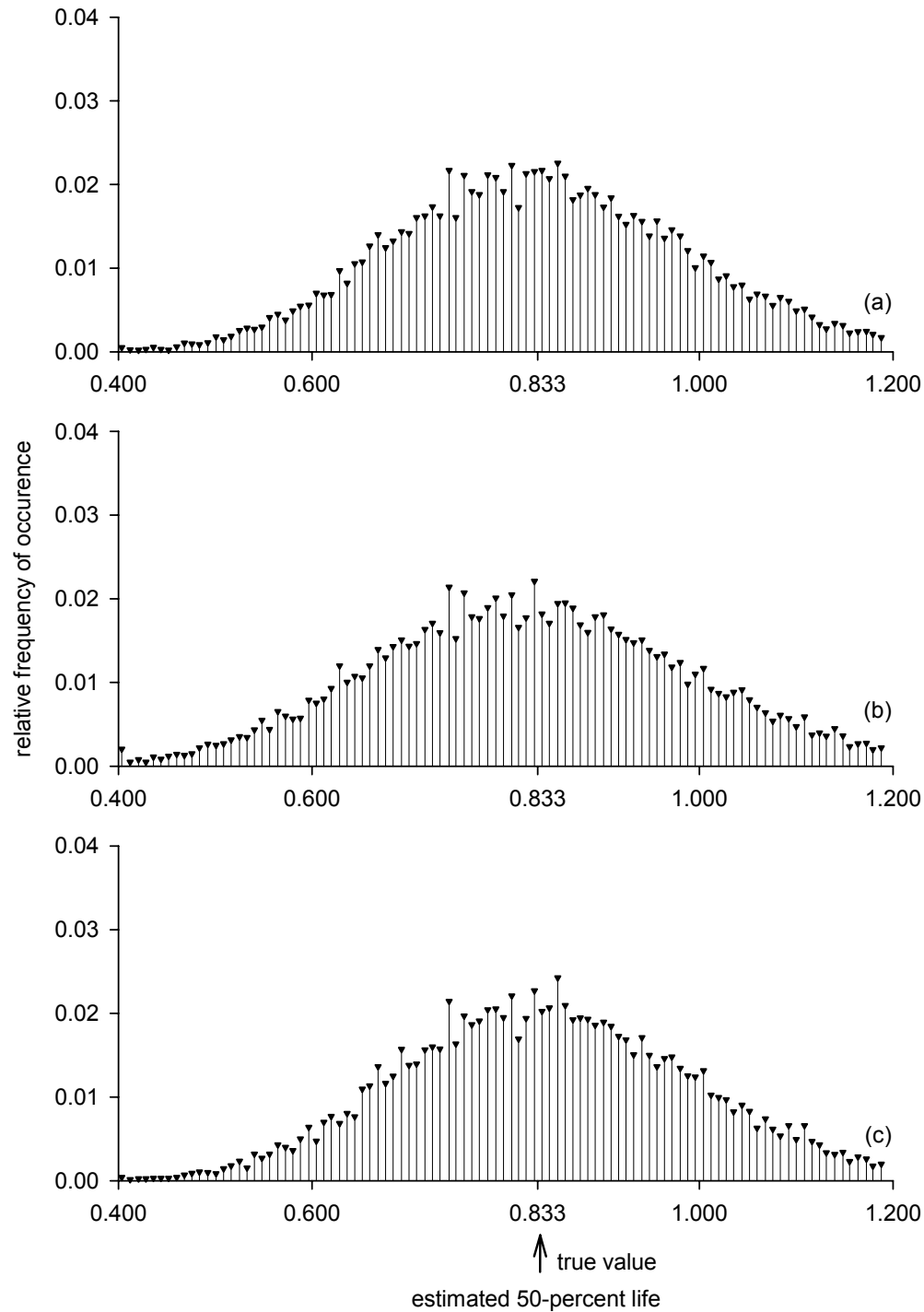


Figure 3.4.8.—Approximate sampling distributions for 50-percent life estimates for the case of 10 random samples selected from a Weibull distribution.
 (a) 2-parameter regression method. (b) 3-parameter regression method.
 (c) Maximum likelihood method.

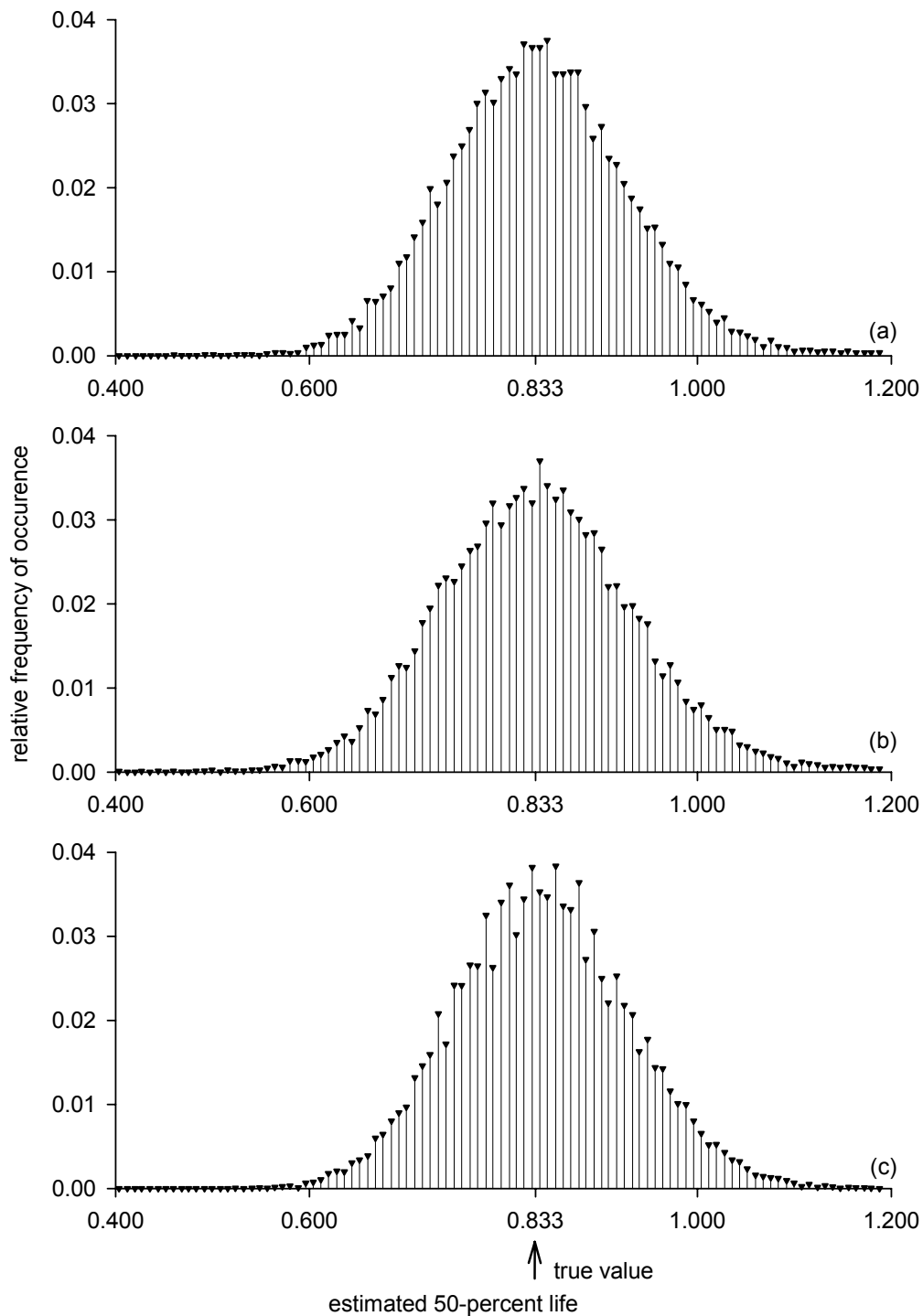


Figure 3.4.9.—Approximate sampling distributions for 50-percent life estimates for the case of 30 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method. (c) Maximum likelihood method.

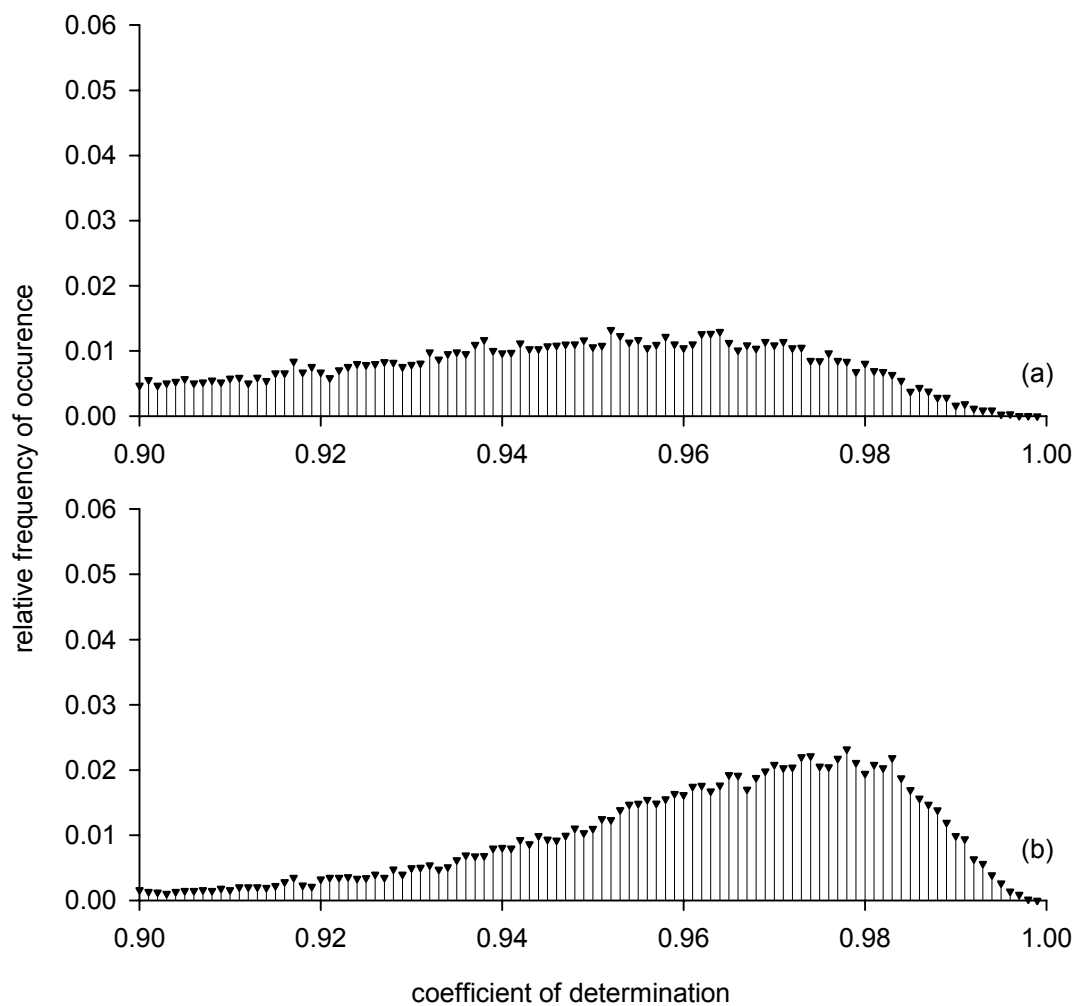


Figure 3.4.10.—Approximate sampling distributions for the coefficient of determination statistic of the regression for the case of 10 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method.

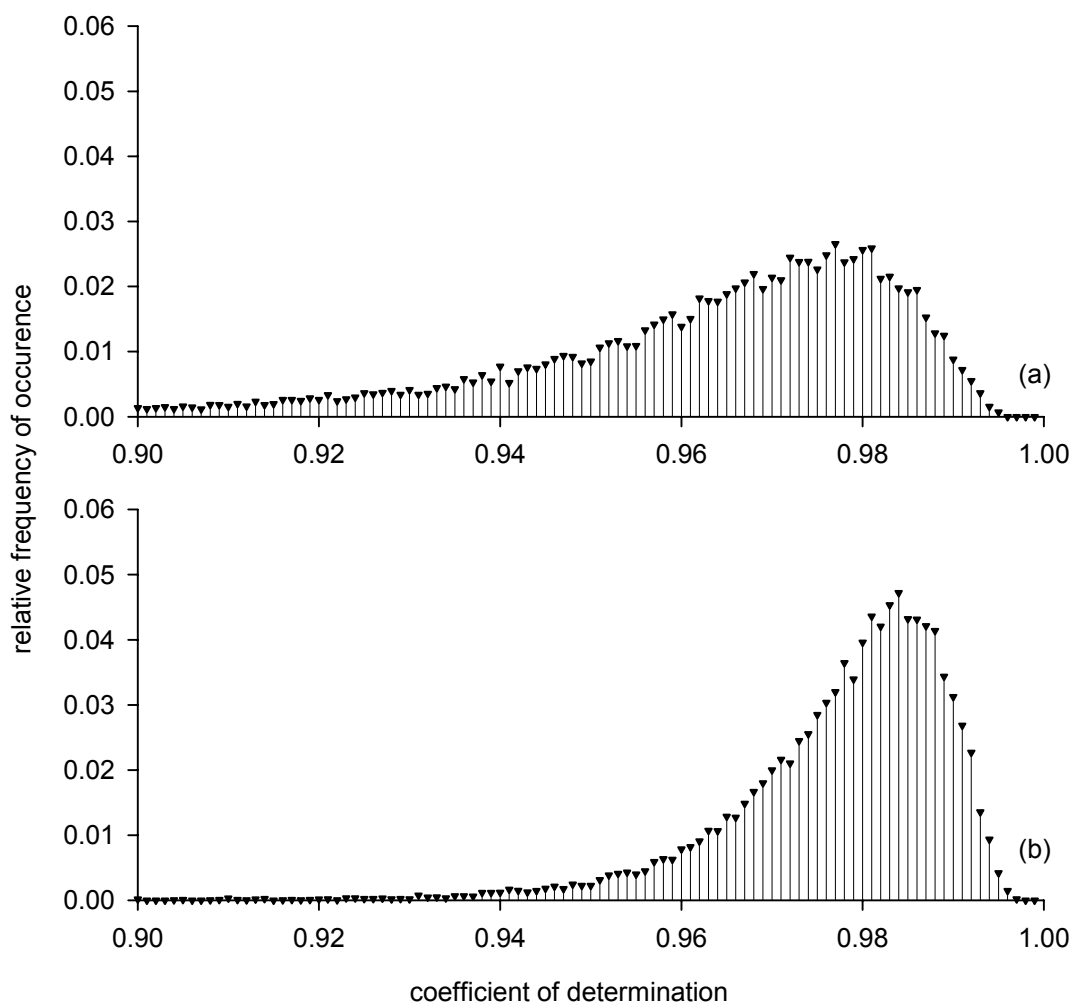


Figure 3.4.11.—Approximate sampling distributions for the coefficient of determination statistic of the regression for the case of 30 random samples selected from a Weibull distribution. (a) 2-parameter regression method. (b) 3-parameter regression method.

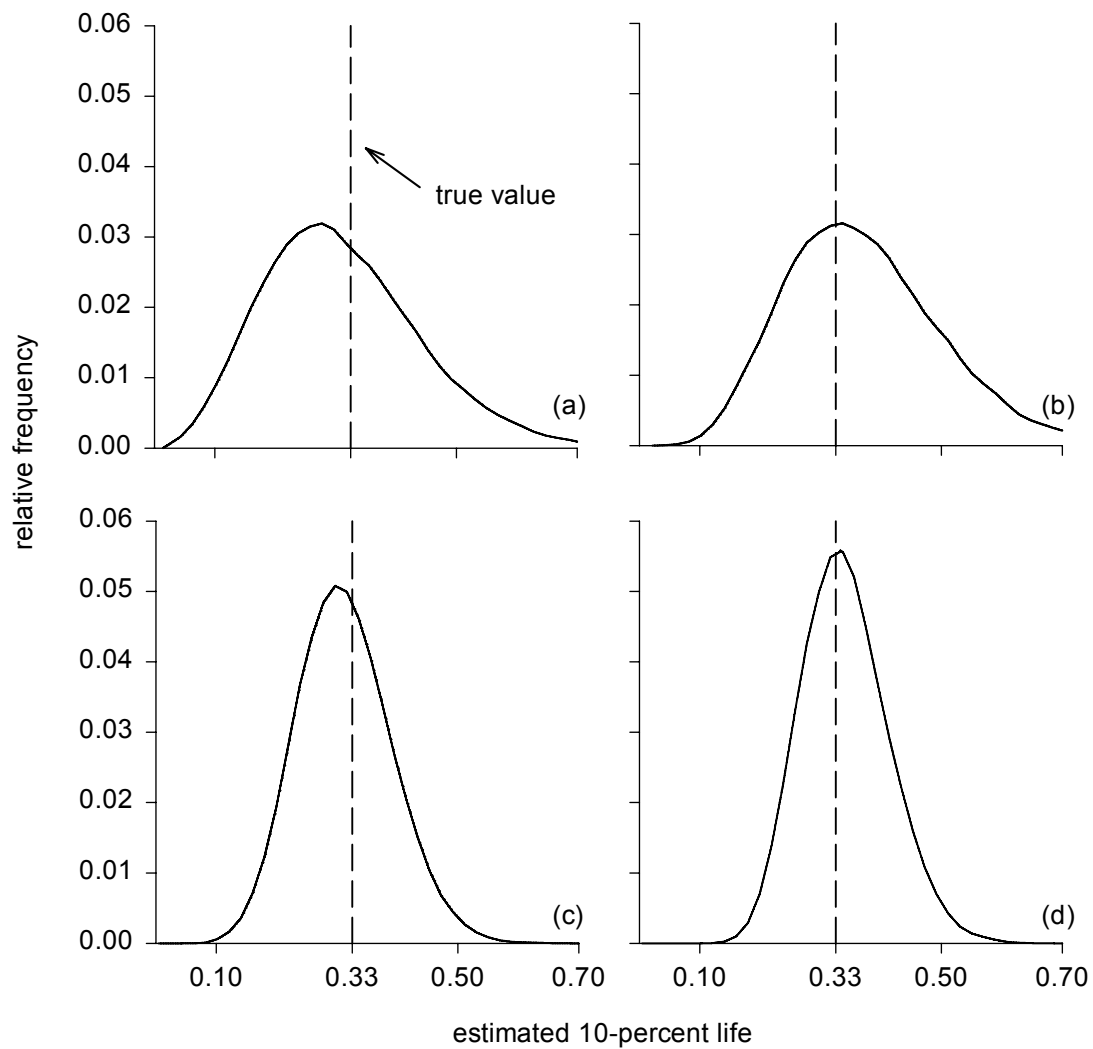


Figure 3.5.1.—Comparison of approximate sampling distributions for 10-percent life estimates highlighting the location of the true value relative to the mode. (a) 2-parameter regression method for sample size of 10. (b) Maximum likelihood method for sample size of 10. (c) 2-parameter regression method for sample size of 30. (d) Maximum likelihood method for sample size of 30.

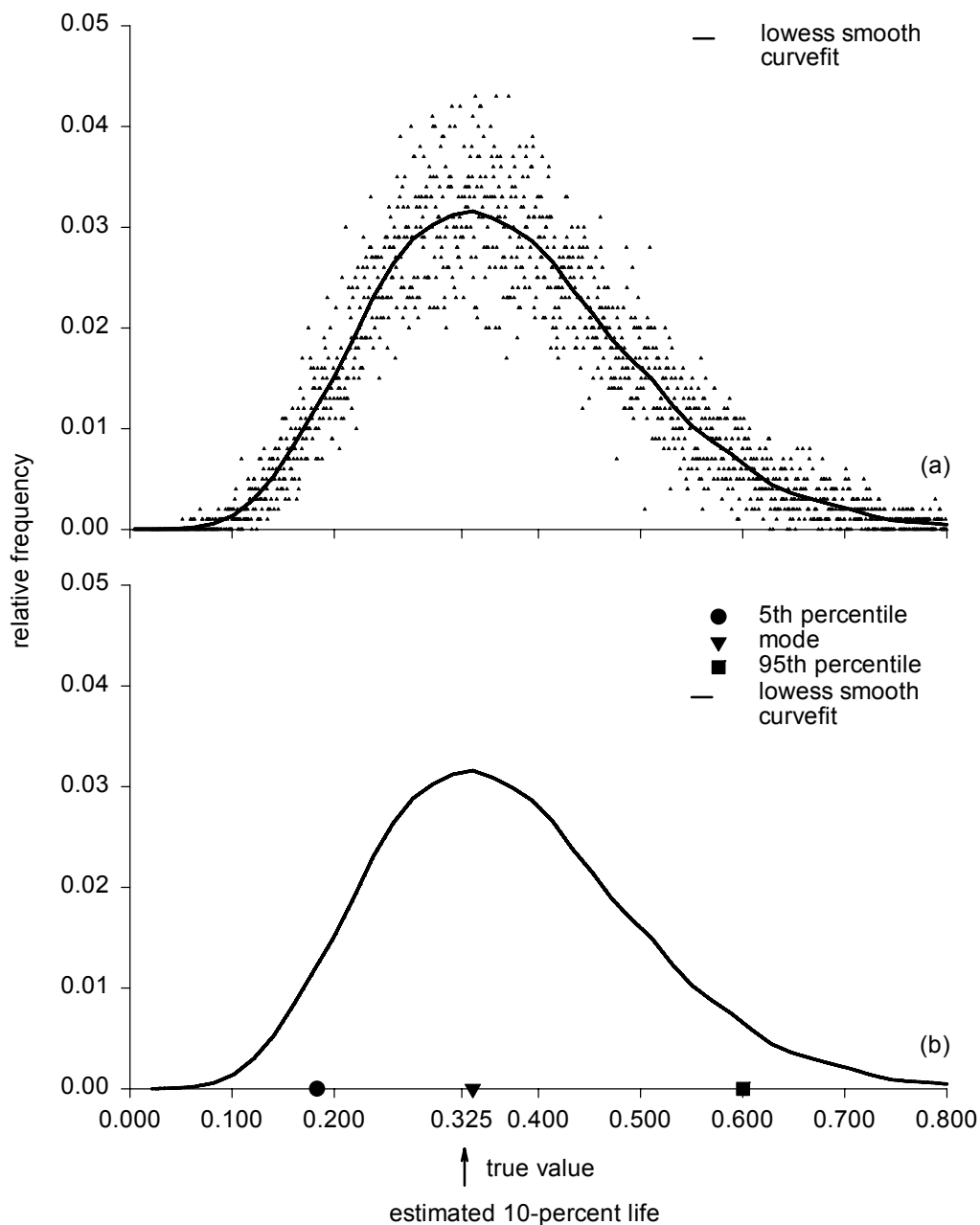


Figure 3.5.2.—Approximate sampling distribution of 10-percent life estimates by fitting a 2-parameter Weibull distribution using the maximum likelihood method. The distribution shown is for the case of 10 samples, randomly selected, from a Weibull distribution with true parameters of threshold = 0.0, scale = 1.0, and shape = 2.0. (a) Histogram as data points with lowess smooth curve. (b) Lowess smooth curve with locations of mode and 90-percent interval endpoints.

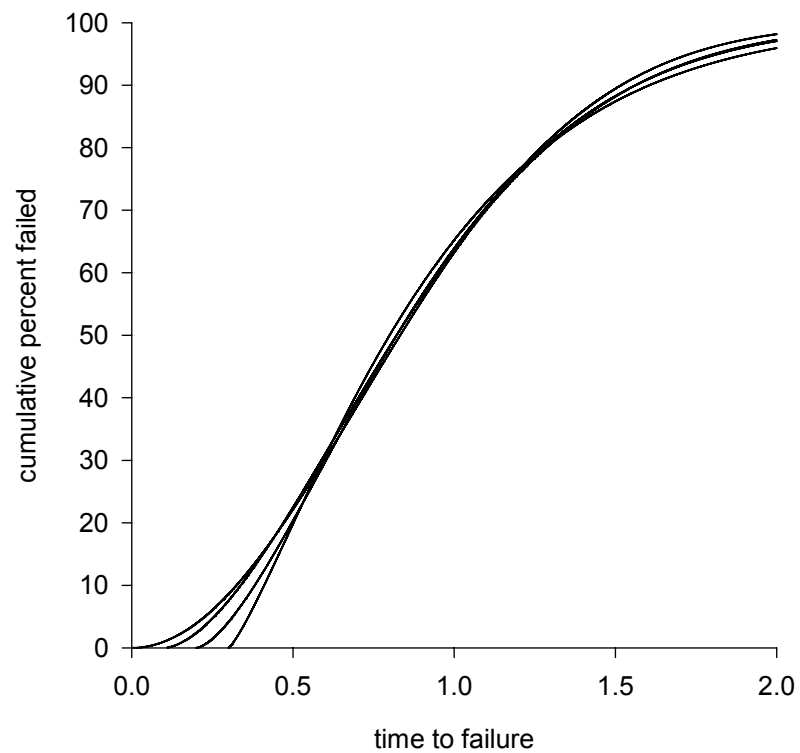


Figure 3.6.1.—Four Weibull distributions with similar lives within the range 10-percent cumulative failed through 90-percent cumulative failed but differing values for the threshold parameters. Defining parameters are provided in table 3.6.1.

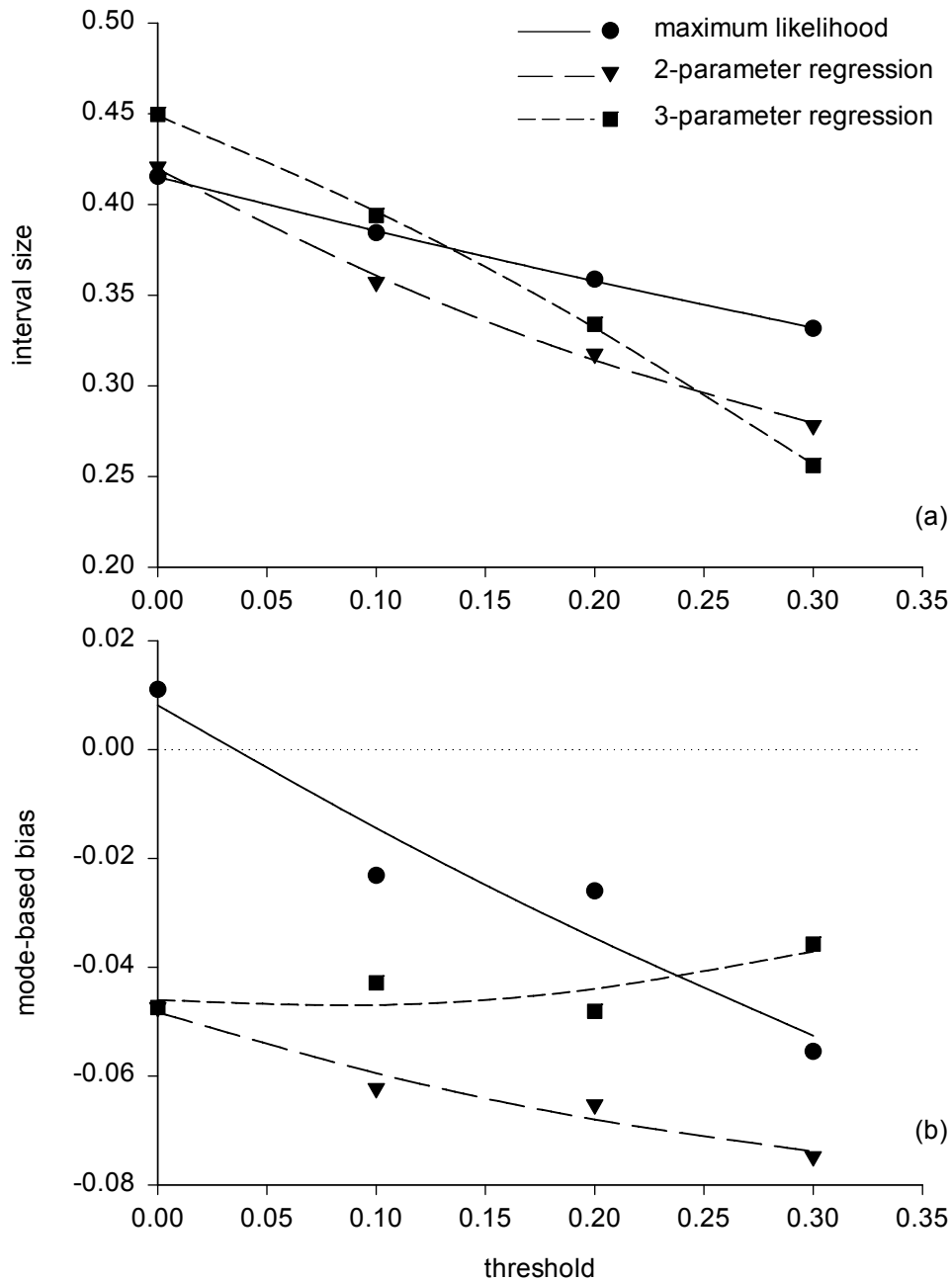


Figure 3.6.2.—Influence of the true threshold value on the performance of three methods for estimating the 10-percent lives as determined by Monte Carlo simulation studies. Ten random samples were available for each estimate. (a) Precision as measured by size of interval bound by 5th and 95th percentiles of ordered estimates. (b) Accuracy as measured by mode-based bias.

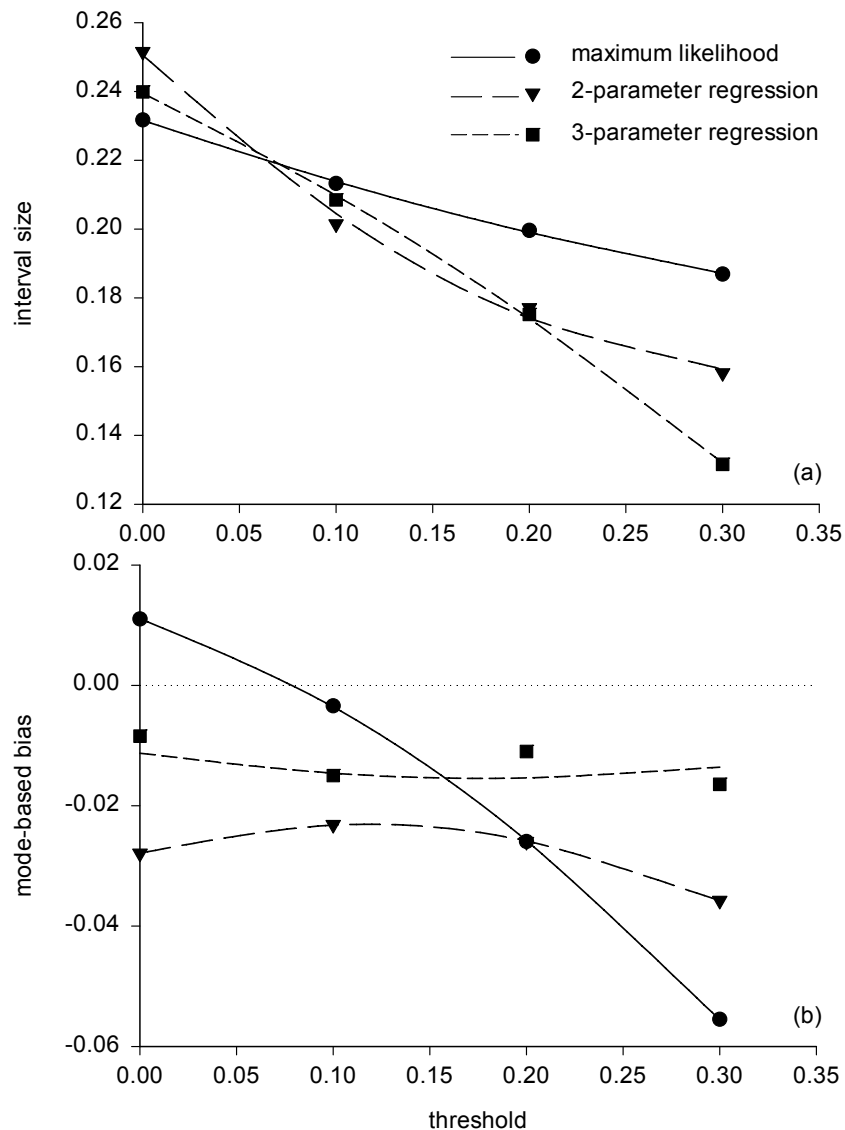


Figure 3.6.3.—Influence of the true threshold value on the performance of three methods for estimating the 10-percent lives as determined by Monte Carlo simulation studies. Thirty random samples were available for each estimate. (a) Precision as measured by size of interval bound by 5th and 95th percentiles of ordered estimates. (b) Accuracy as measured by mode-based bias.

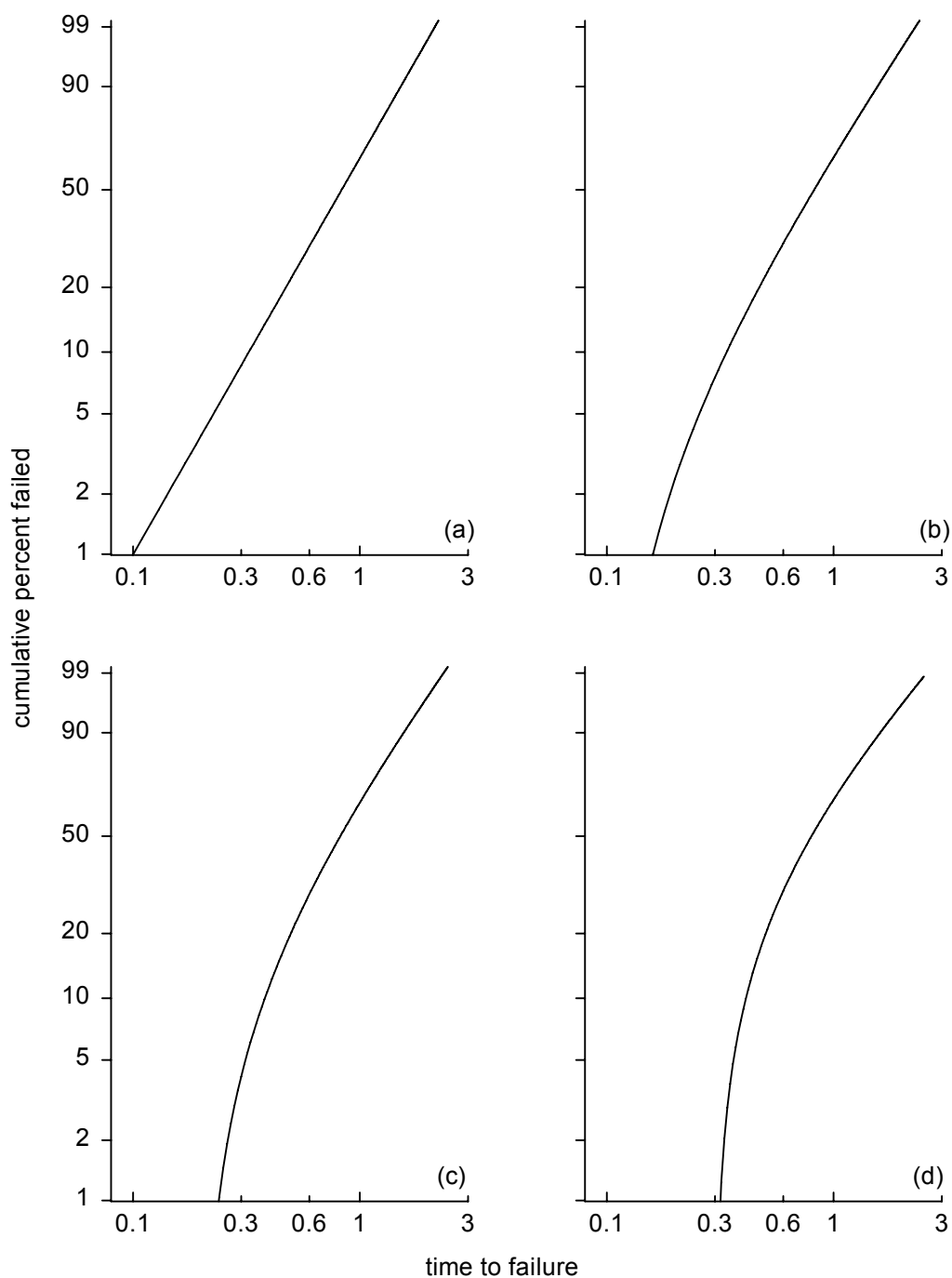


Figure 3.6.4.—Weibull plots of four Weibull distributions with similar lives within the range 10-percent cumulative failed through 90-percent cumulative failed but differing values for the threshold parameters. Defining parameter are provided in table 3.6.1.

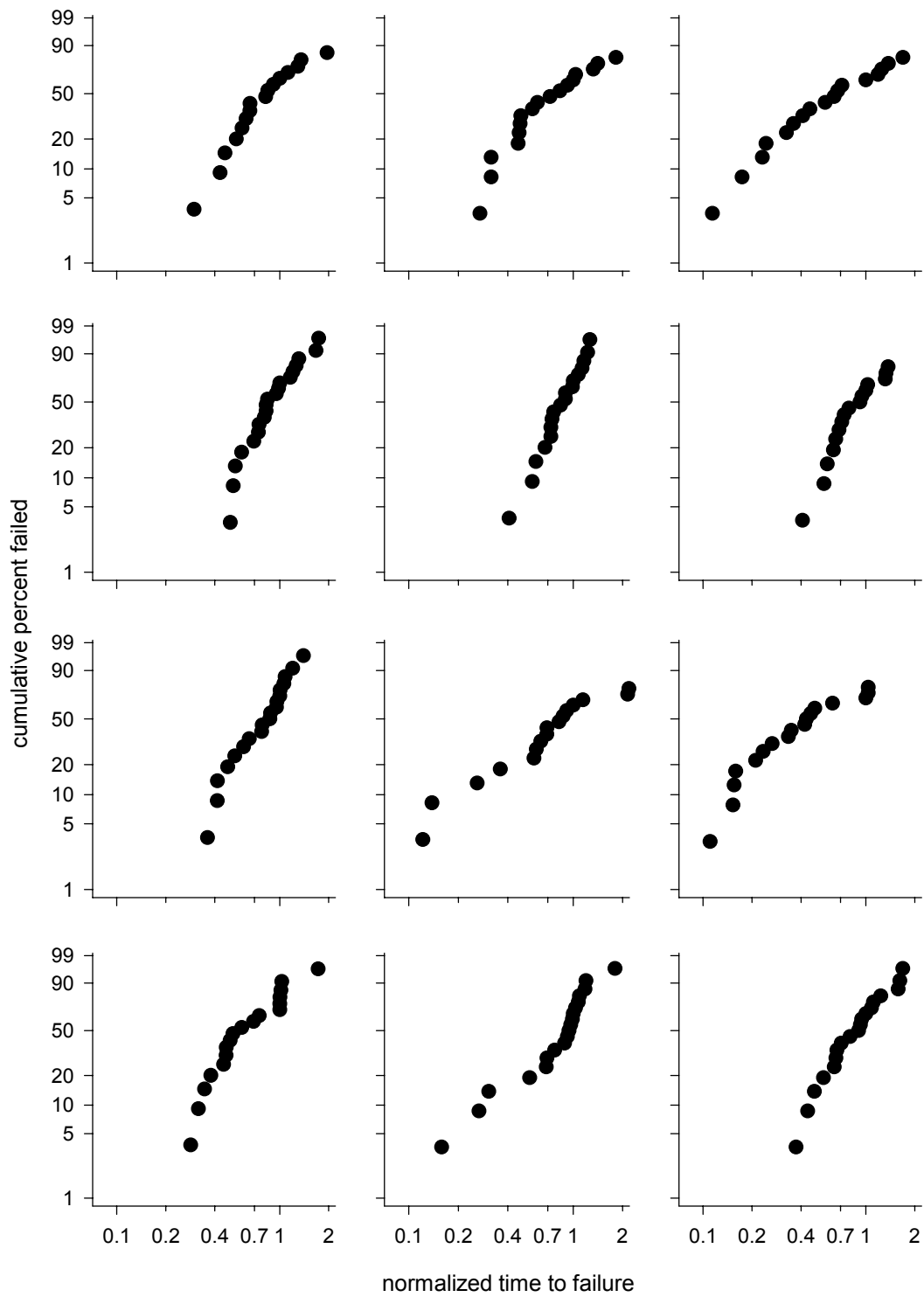


Figure 3.6.5.—Twelve sets of gear fatigue displayed as Weibull plots using a normalized time to failure.

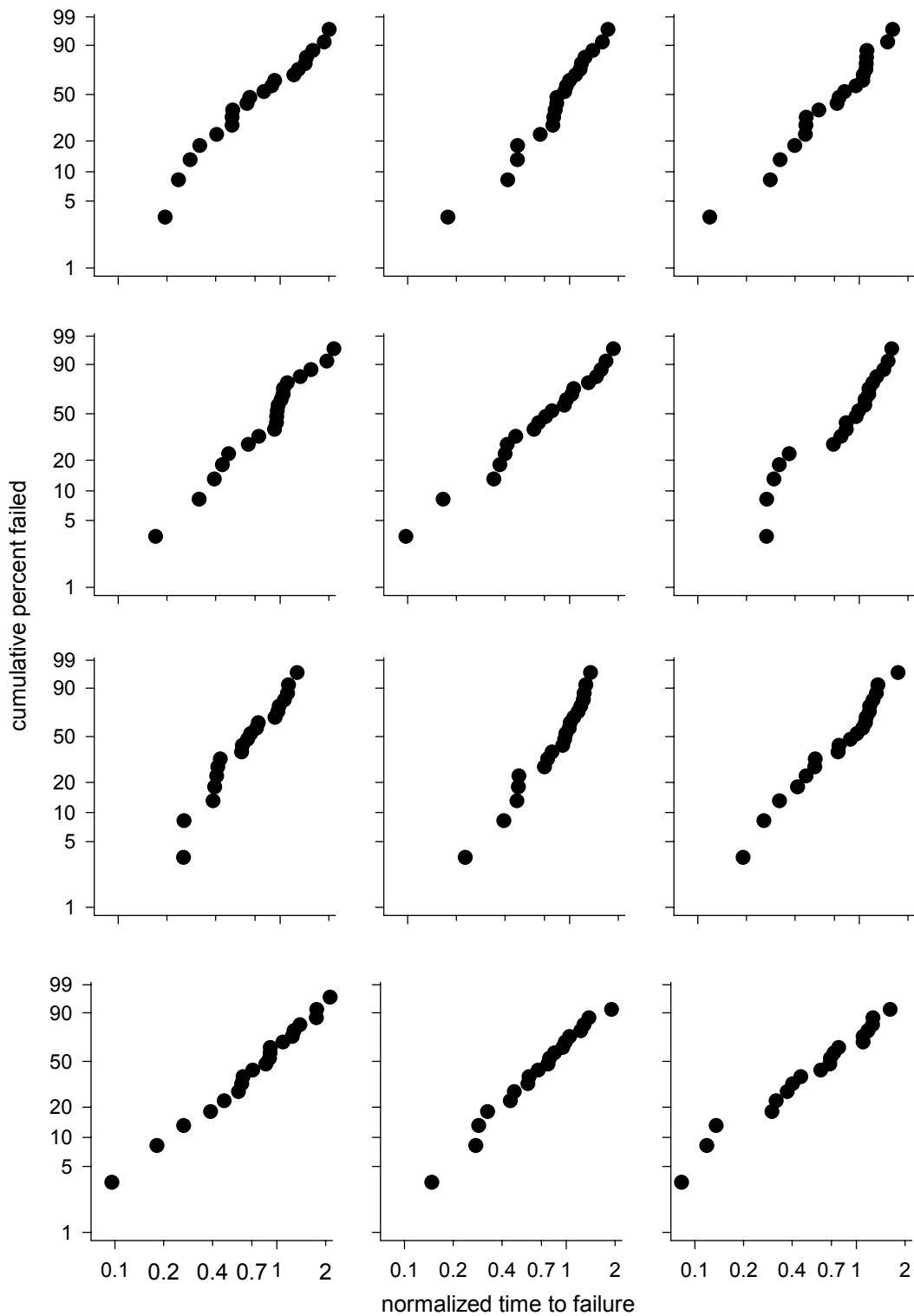


Figure 3.6.6.—Weibull plots containing 20 psuedo-random samples taken from a Weibull distribution with defining parameters of scale = 1.0, shape = 2.0, and threshold = 0.0.

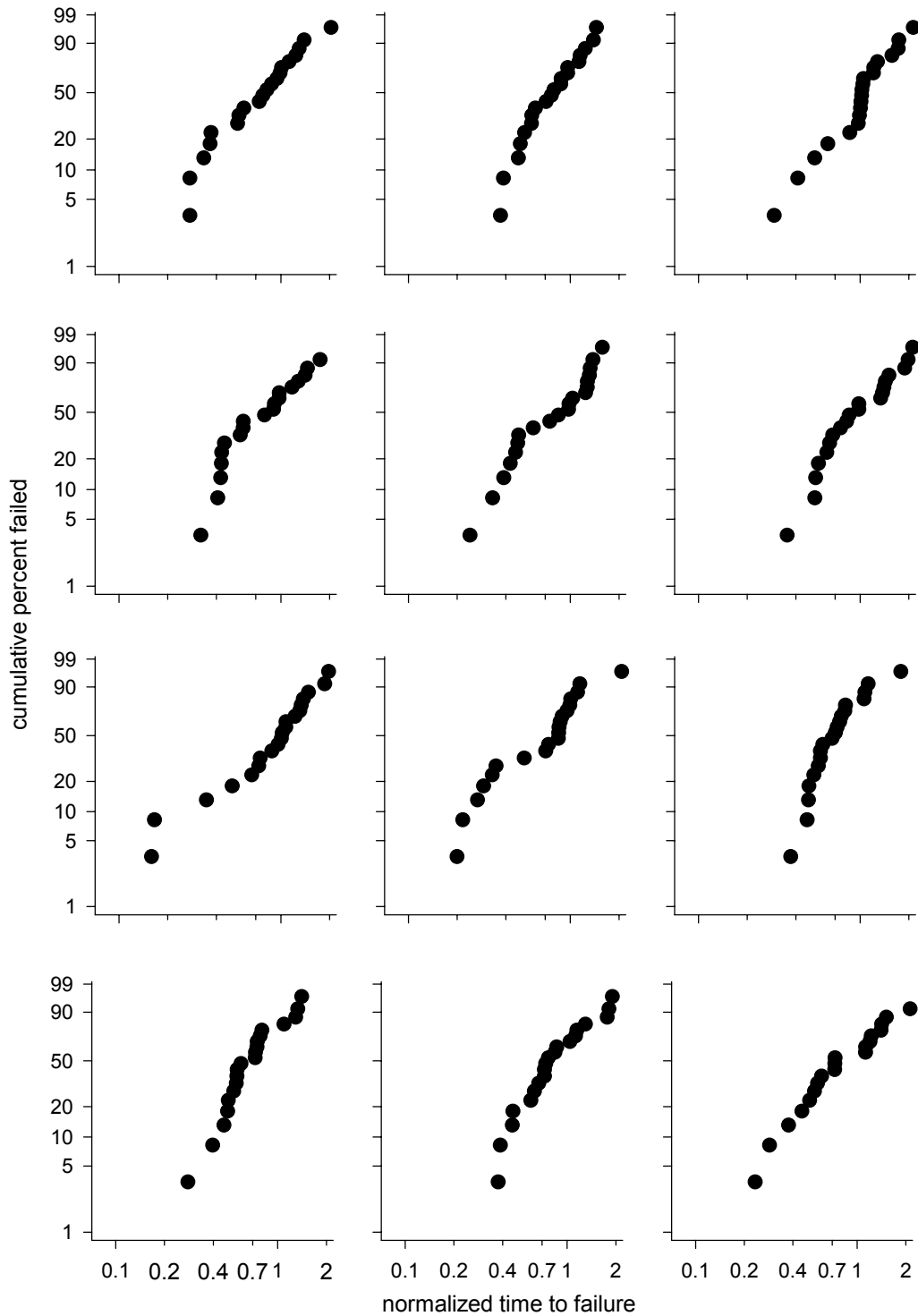


Figure 3.6.7.—Weibull plots containing 20 psuedo-random samples taken from a Weibull distribution with defining parameters of scale = 0.90, shape = 1.70, and threshold = 0.1.

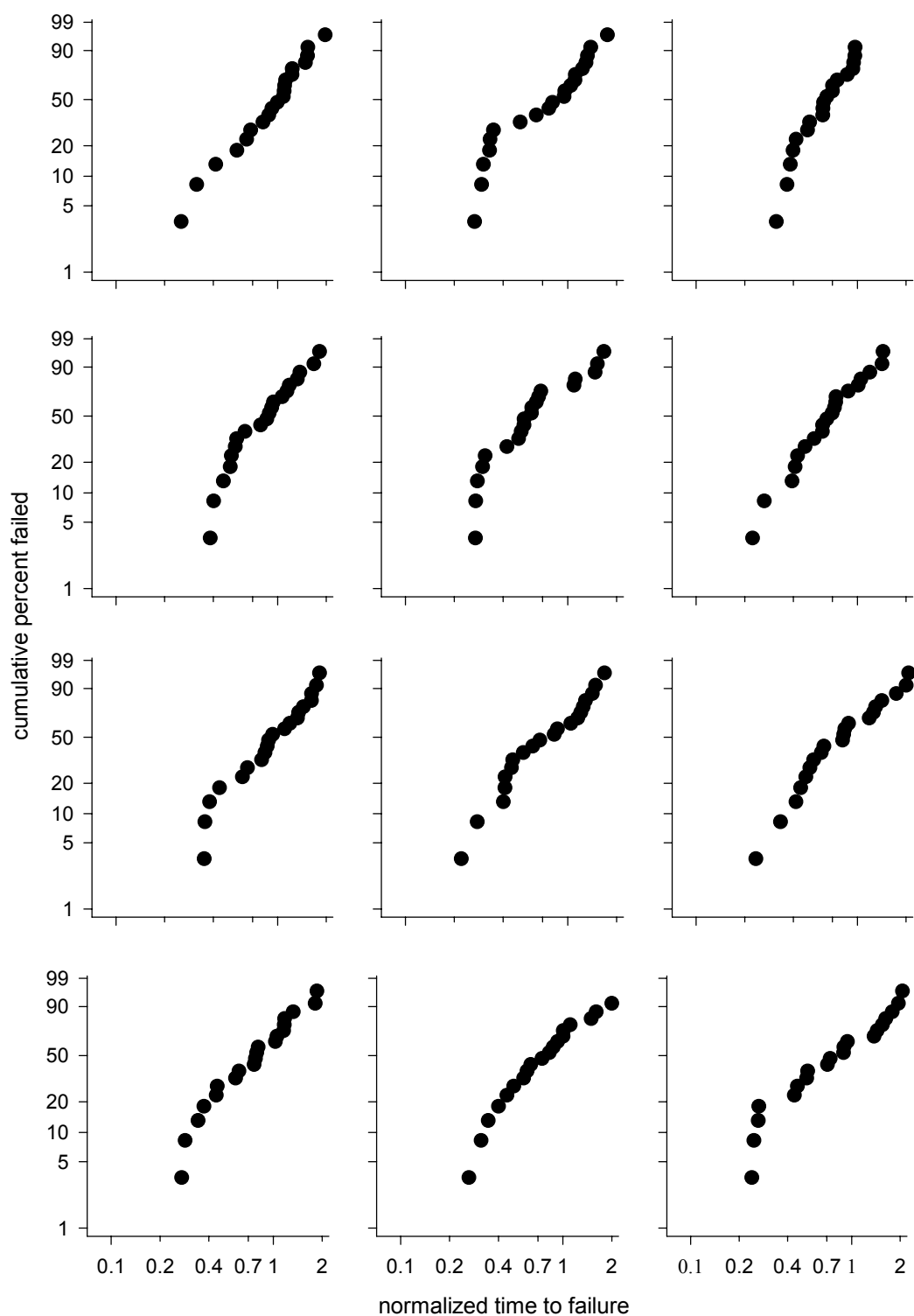


Figure 3.6.8.—Weibull plots containing 20 psuedo-random samples taken from a Weibull distribution with defining parameters of scale = 0.769, shape = 1.445, and threshold = 0.2.

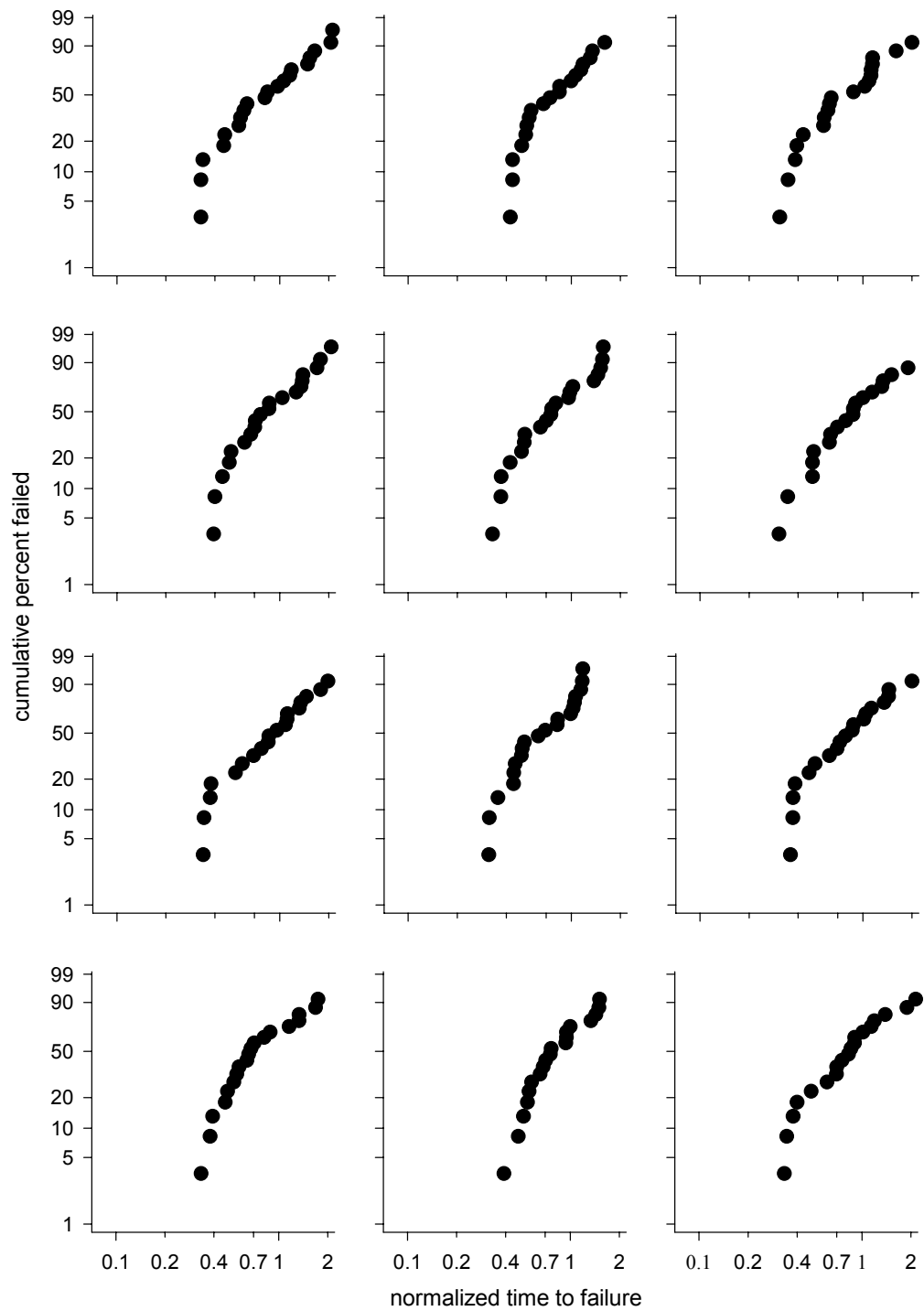


Figure 3.6.9.—Weibull plots containing 20 psuedo-random samples taken from a Weibull distribution with defining parameters of scale = 0.639, shape = 1.145, and threshold = 0.3.

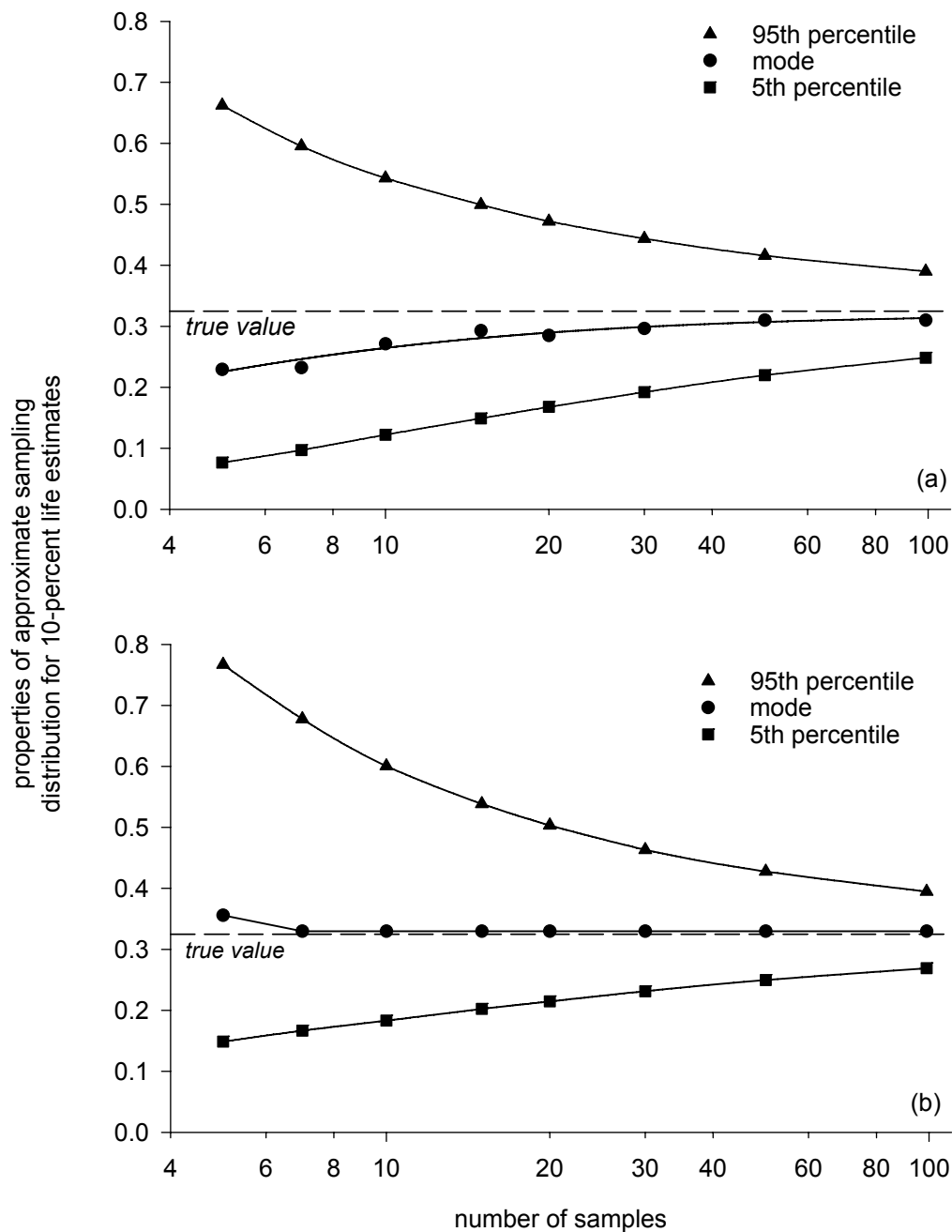


Figure 3.7.1.—Properties of approximate sampling distributions for 10-percent life estimates, as determined by Monte Carlo simulation, as a function of number of samples. The population from which samples were drawn is distributed Weibull with parameters shape = 2.0, scale = 1.0, and threshold = 0.0. (a) Exact median rank least-squares regression method (b) Maximum likelihood method.

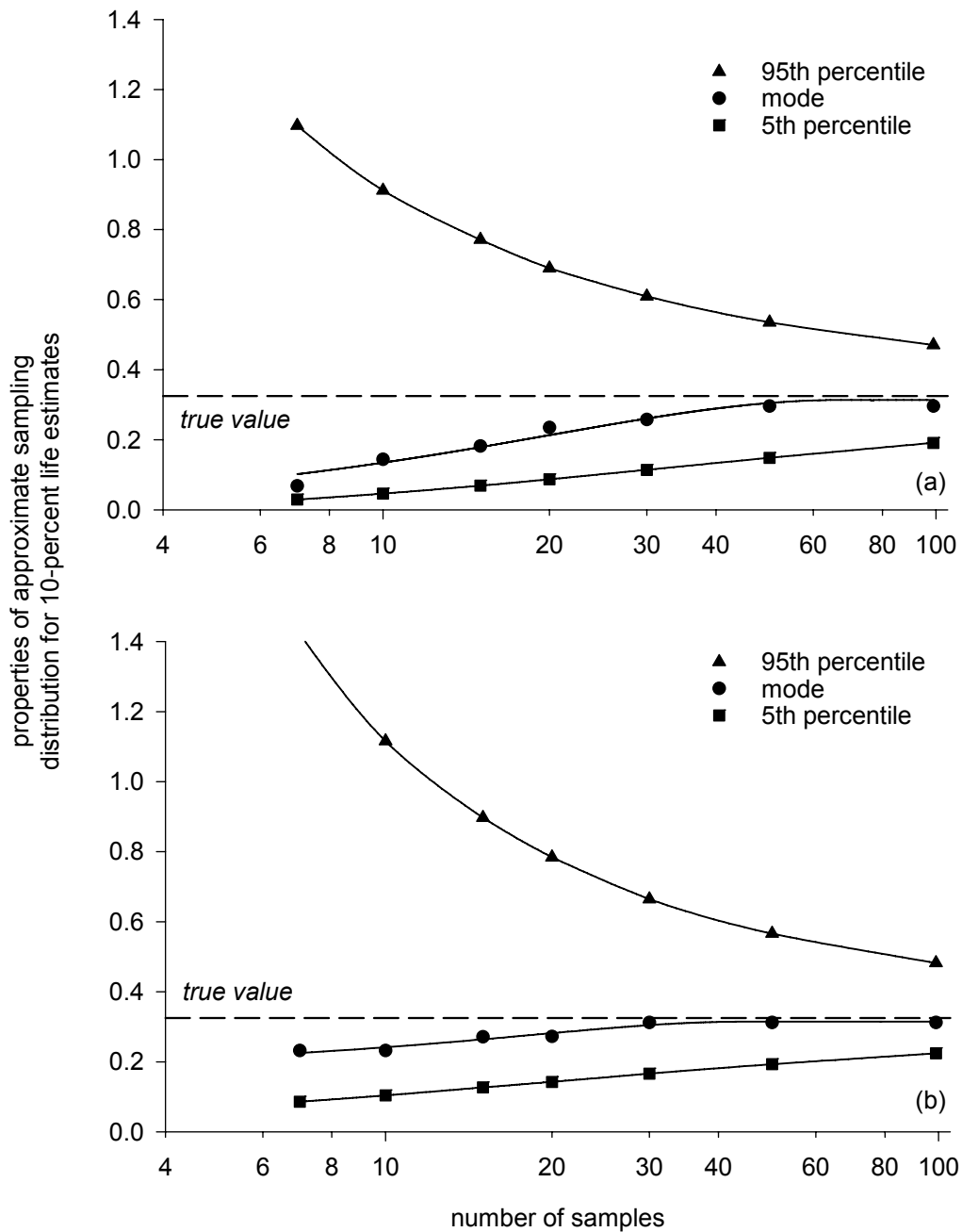


Figure 3.7.2.—Properties of approximate sampling distributions for 10-percent life estimates, as determined by Monte Carlo simulation, as a function of number of samples. The population from which samples were drawn is distributed Weibull with parameters shape = 1.0, scale = 3.09, and threshold = 0.0. (a) Exact median rank least-squares regression method. (b) Maximum likelihood method.

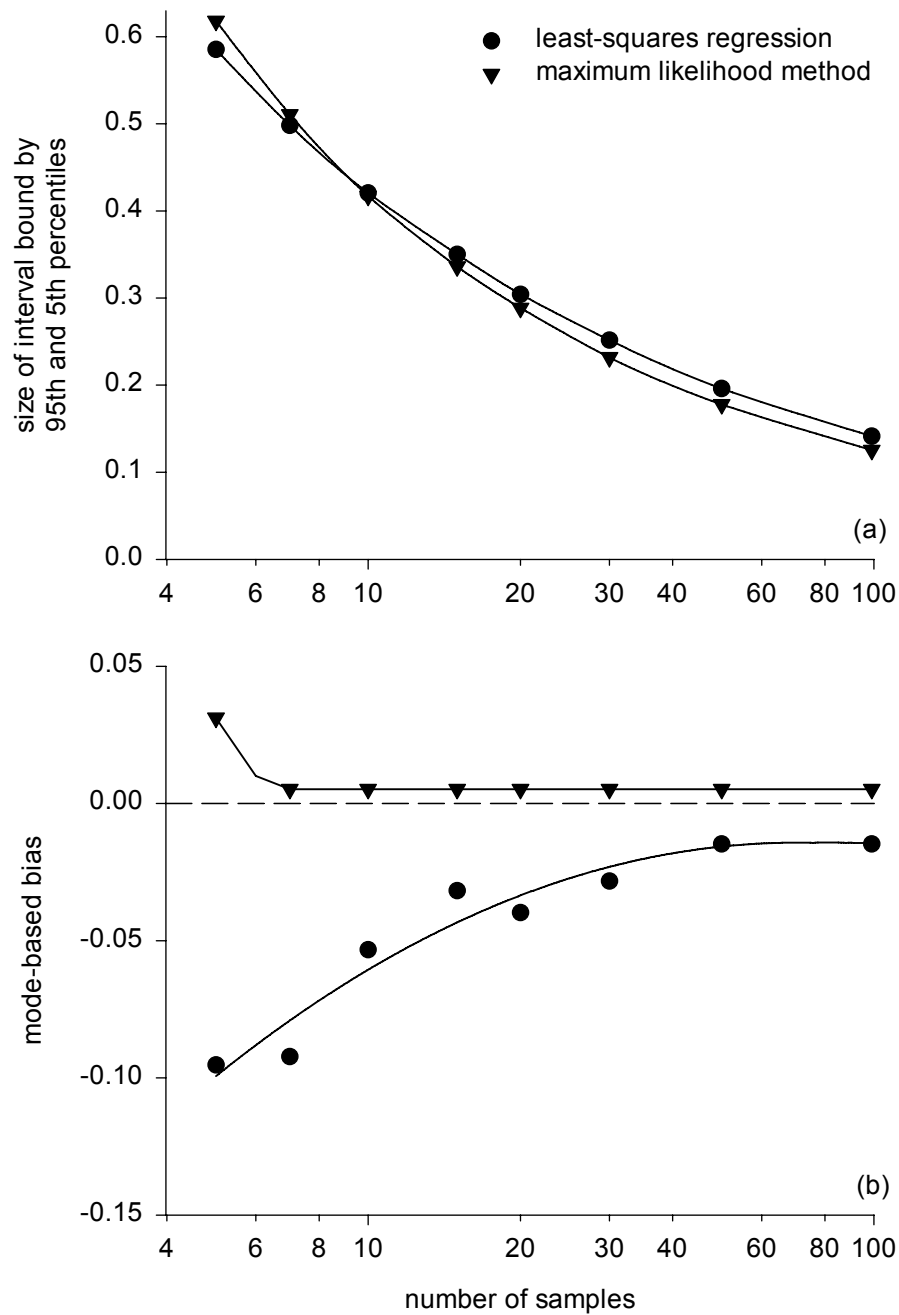


Figure 3.7.3.—Properties of approximate sampling distributions for 10-percentile estimates, as determined by Monte Carlo simulation, as a function of number of samples. The population from which samples were drawn is distributed Weibull with parameters shape = 2.0, scale = 1.0, and threshold = 0.0. (a) Precision of estimates as measured by 90 percent interval. (b) Accuracy of estimates as measured by mode-based bias.

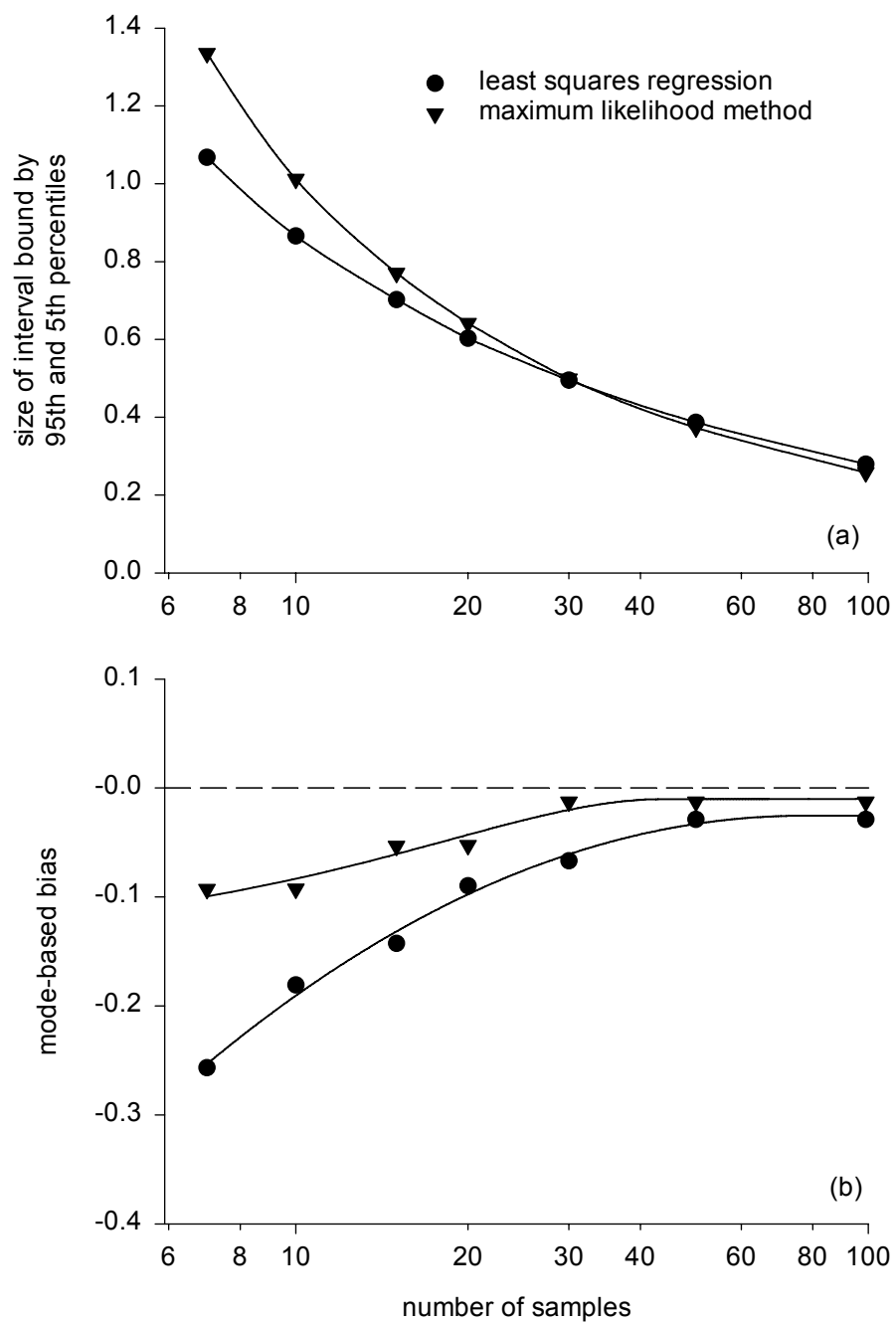


Figure 3.7.4.—Properties of approximate sampling distributions for 10-percent life estimates, as determined by Monte Carlo simulation, as a function of number of samples. The population from which samples were drawn is distributed Weibull with parameters shape = 1.0, scale = 3.09, and threshold = 0.0. (a) Precision of estimates as measured by 90 percent interval (b) Accuracy of estimates as measured by mode-based bias.

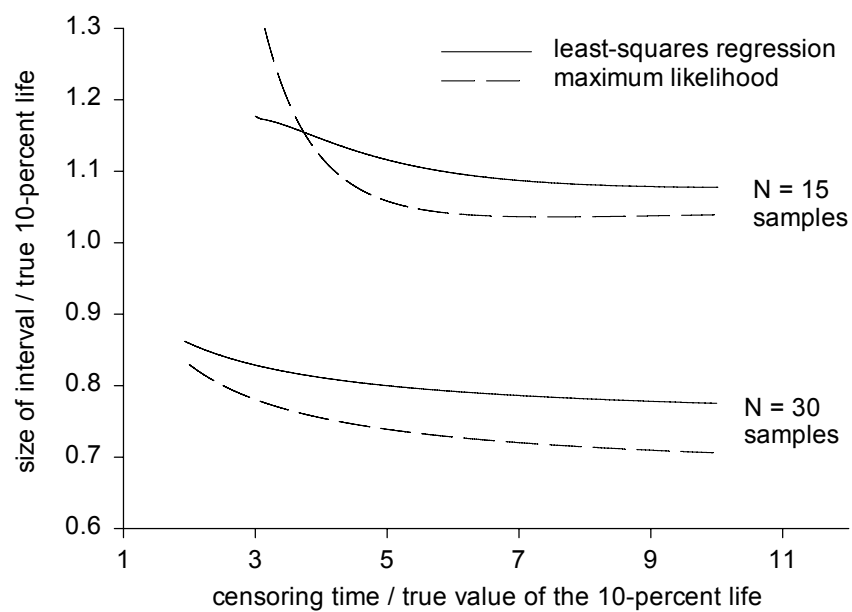


Figure 3.8.1.—The influence of censoring time on the size of the interval bound by the 5th and 95th percentiles of the ordered estimates for the 10-percent life. Data shown is for the case of random samples taken from distribution with shape parameter = 2.0, scale parameter = 1.0, threshold parameter = 0.0.

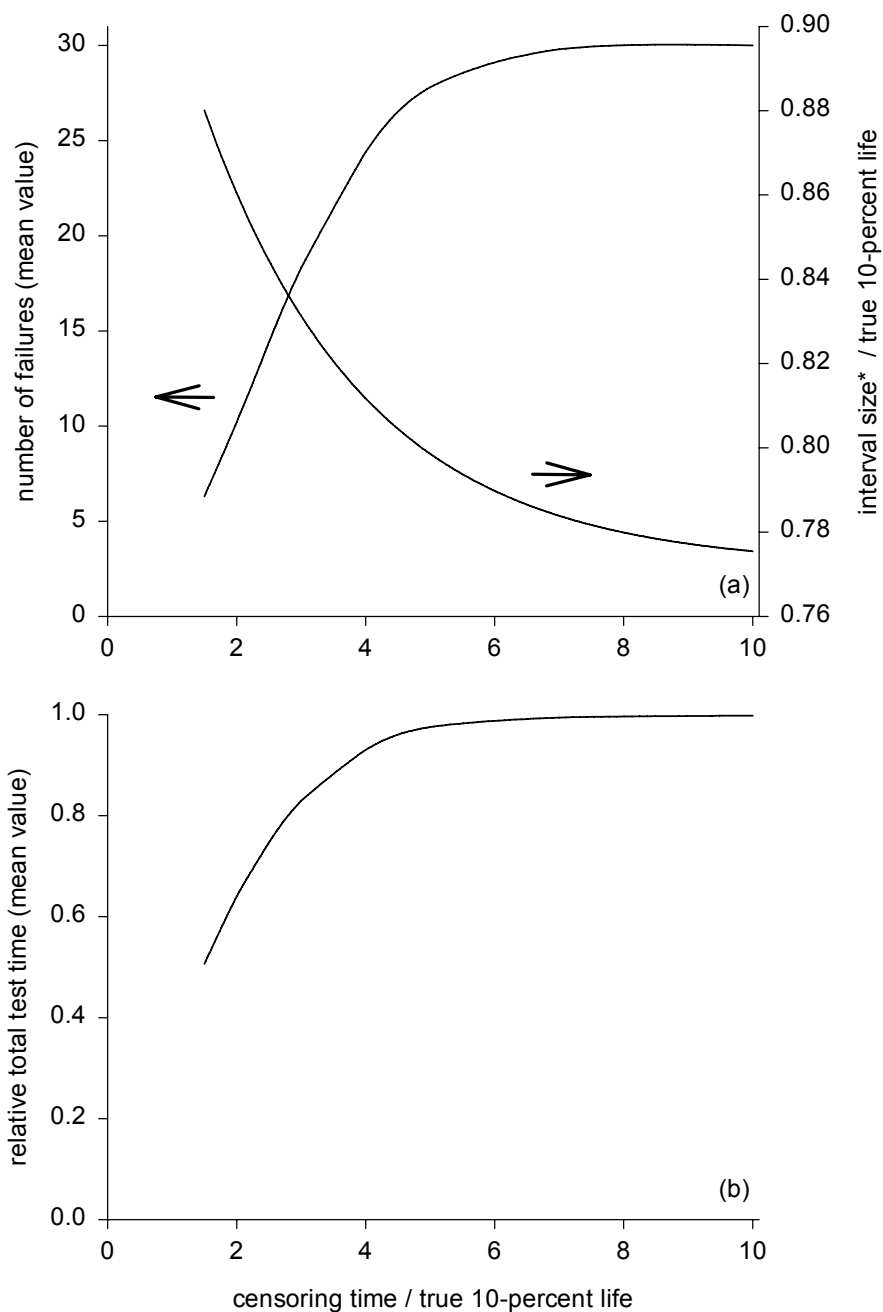


Figure 3.8.2.—The influence of censoring time on the breadth of the sampling distribution, mean value of the number of failures, and mean value of the relative total test time. Data shown is for the case of 30 random samples taken from distribution with shape parameter = 2.0, scale parameter = 1.0, threshold parameter = 0.0, and analysis by the maximum likelihood method. (a) Interval size and number of failures. (b) Relative total test time. *Interval size is that bounded by the 5th and 95th percentiles of the sorted estimates

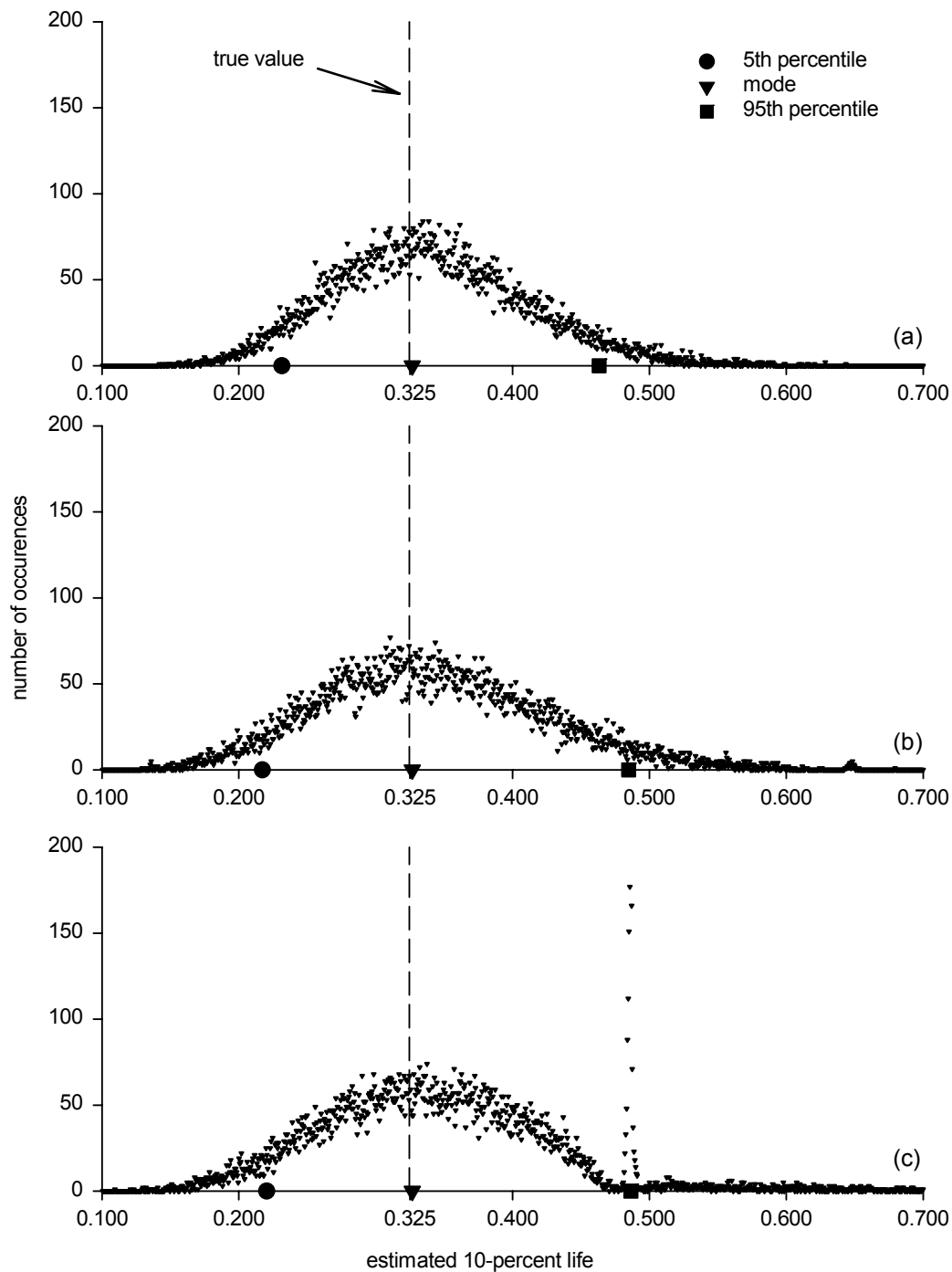


Figure 3.8.3.—The influence of censoring time on the sampling distribution of estimated 10-percent life. Data shown is for the case of 30 random samples taken from distribution with shape parameter = 2.0, scale parameter = 1.0, threshold parameter = 0.0 and analysis by the maximum likelihood method. (a) Censoring time = 3.2. (b) Censoring time = 0.65. (c) Censoring time = 0.49.

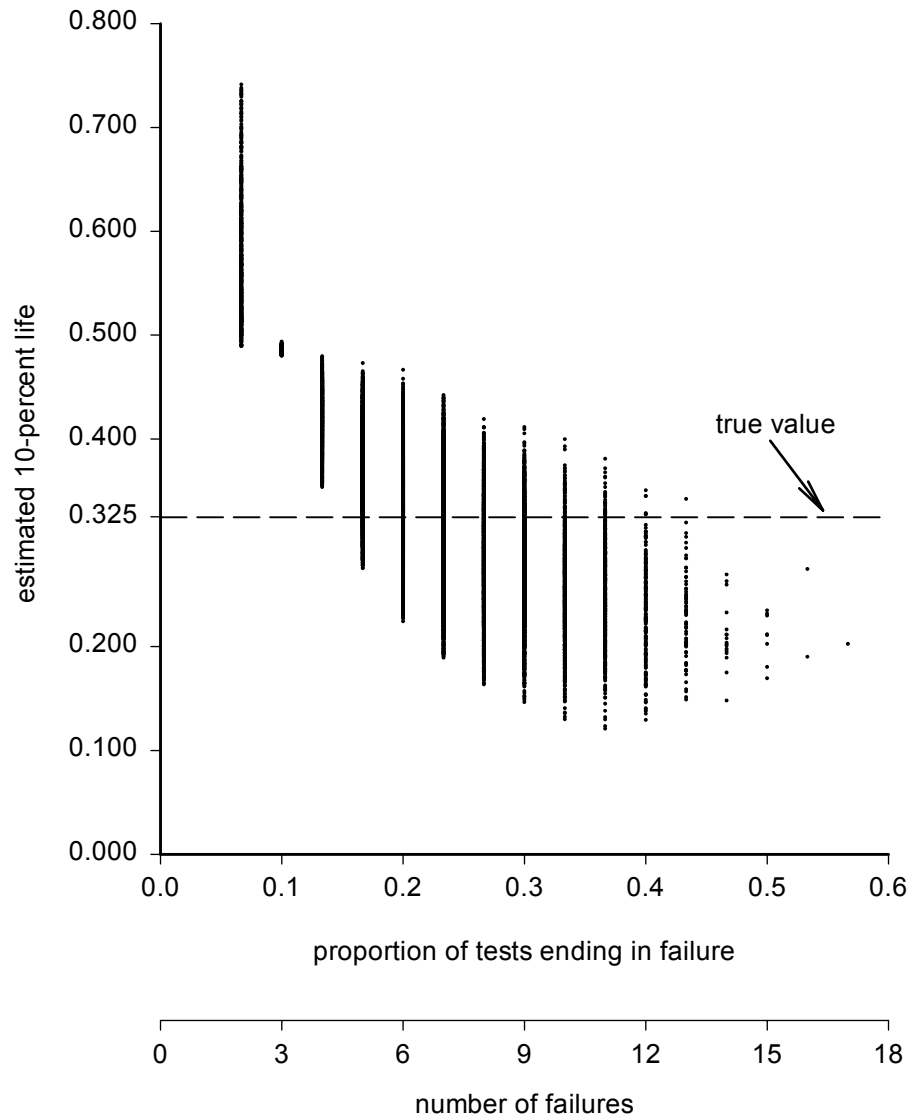


Figure 3.8.4.—The influence of the proportion of tests ending in failures on the ranges for the estimated 10-percent life. The data shown is for the case of 30 random samples taken from a Weibull distribution with shape parameter = 2.0, scale parameter = 1.0, and threshold parameter = 0.0. The simulated censoring time used = 0.49, and analysis was by the maximum likelihood method.

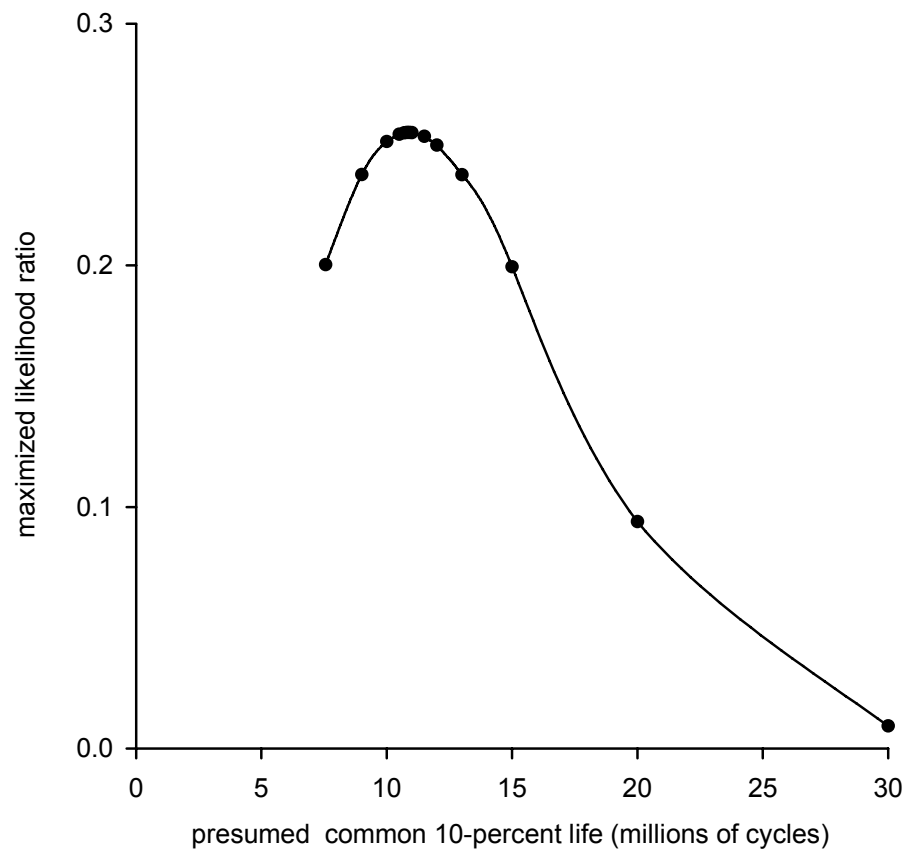


Figure 3.11.1.—Maximized likelihood ratio as a function of the presumed common 10-percent life. The common 10-percent life providing the largest likelihood ratio is the value 10.9.

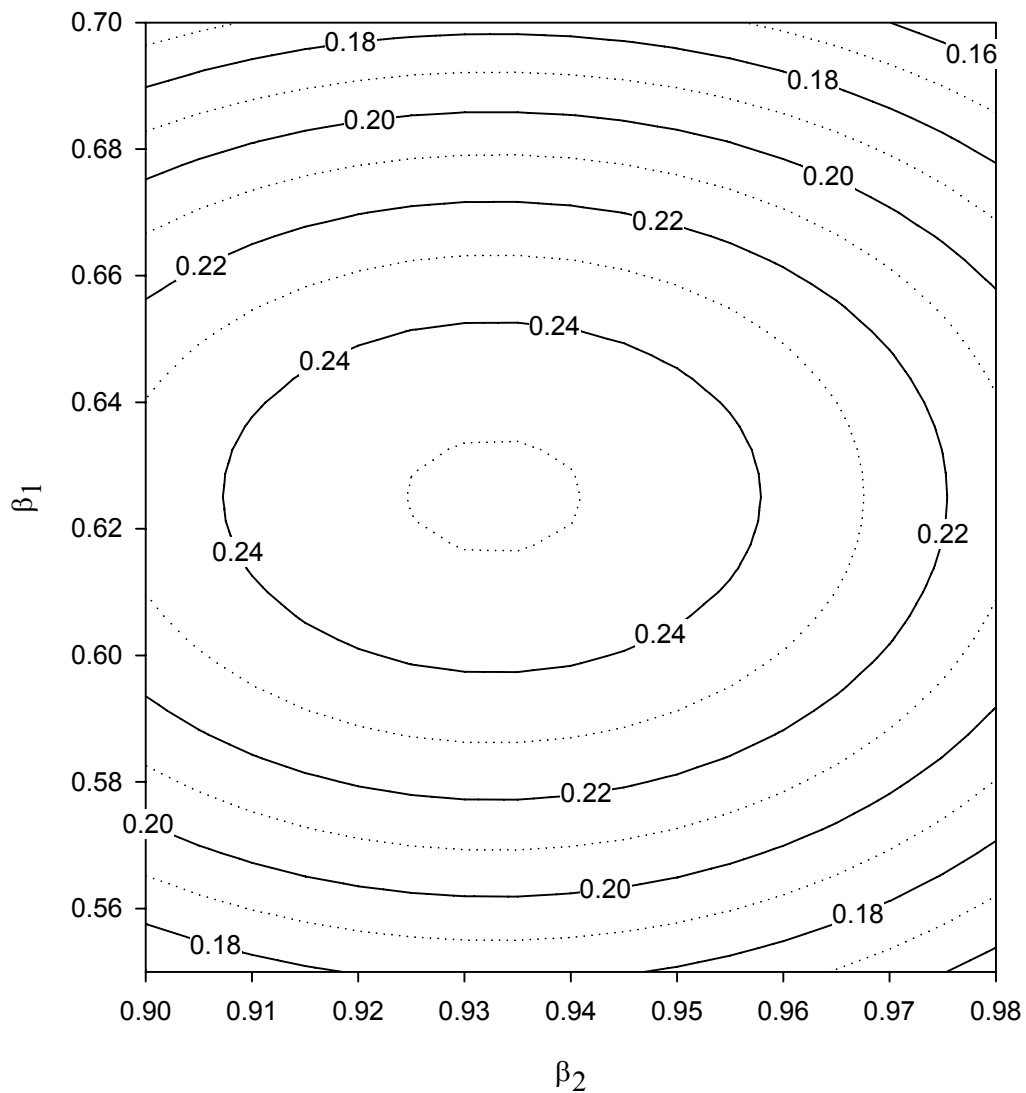


Figure 3.11.2.—Contours of equal likelihood ratios as a function of shape parameter values for the case of a proposed common 10-percent life of 10.0. The maximum ratio is 0.2512 for the parameter values $\beta_1 = 0.625$ and $\beta_2 = 0.932$.

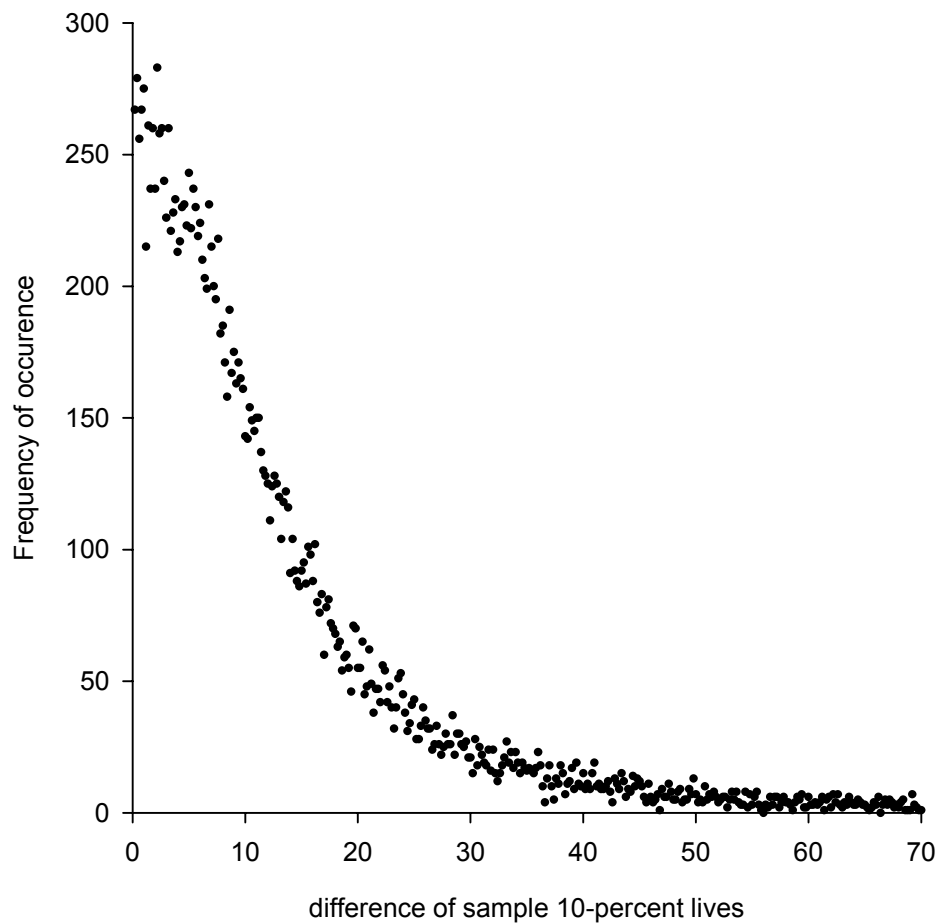


Figure 3.11.3.—Approximate sampling distribution for the difference of sample 10-percent lives as produced by Monte Carlo simulation.

Chapter 4—Experimental Evaluation of Gear Surface Fatigue Life

4.1 Introduction

The power density of a gearbox is an important consideration for many applications and is especially important for gearboxes used on aircraft. One factor that limits gearbox power density is the ability of the gear teeth to transmit power for the required number of cycles without pitting or spalling. Economical methods for improving surface fatigue lives of gears are therefore highly desirable.

Tests of rolling element bearings (refs. 4.1 to 4.4, for example) have shown that the bearing life is affected by the lubricant viscosity and the surface roughness. When the specific film thickness (the lubricant film thickness divided by the composite surface roughness) is less than unity, the service life of the bearing is considerably reduced. The effect of oil viscosity and surface finish on the scoring load capacity of gears was investigated experimentally more than 40 years ago (ref. 4.5). More recently, some investigators have anticipated that the effect of specific film thickness on gear life could be even more pronounced than the effect on bearing life (ref. 4.6). To improve the surface fatigue lives of gears, the film thickness may be increased, the composite surface roughness reduced, or both approaches may be adopted. These two effects have been studied.

Townsend and Shimski (ref. 4.7) studied the influence of seven different lubricants of varying viscosity on gear fatigue lives. Tests were conducted on a set of case-carburized and ground gears, all manufactured from the same melt of consumable-electrode vacuum-melted (CVM) AISI 9310 steel. At least 17 pairs of surfaces were tested with each lubricant. They noted a strong positive correlation of the gear surface fatigue lives with the calculated film thickness and demonstrated that increasing the film thickness does indeed improve gear surface fatigue life.

At least three investigations have been carried out to demonstrate the relation between gear surface fatigue and surface roughness. One investigation by Tanka, et al. (ref. 4.8) involved a series of tests conducted on steels of various chemistry, hardness, and states of surface finish. Some gears were provided with a near-mirror finish by using a special grinding wheel and machine (ref. 4.9). The grinding procedure was a generating process that provided teeth with surface roughness quantified as R_{\max} of about $0.1\text{ }\mu\text{m}$ ($4\text{ }\mu\text{in.}$). A series of pitting durability tests were conducted and included tests of case-carburized pinions mating with both plain carbon steel gears and through-hardened steel gears. They concluded that the gear surface durability was improved in all cases because of the near-mirror finish. They noted that when a case-hardened, mirror-finished pinion was mated with a relatively soft gear, the gear became polished with running. They considered that this polishing during running improved the surface durability of the gear. None of the tests conducted in the study, however, included a case-carburized pinion mated with a case-carburized gear.

A second investigation by Nakasuji et al. (refs. 4.10 and 4.11) studied the possibility of improving gear fatigue lives by electrolytic polishing. They conducted their tests using medium carbon steel gears and noted that the electropolishing process altered the gear profile and the surface hardness as well as the surface roughness. The polishing reduced the surface hardness and changed the tooth profiles to the extent that the measured dynamic tooth stresses were significantly larger relative to the ground gears. Even though the loss of hardness and increased dynamic stresses would tend to reduce stress limits for pitting durability, the electrolytic polishing was shown to improve the stress limit, at which the gears were free of pitting, by about 50 percent.

Hoyashita et al. (refs. 4.12 and 4.13) completed a third investigation of the relation between surface durability and roughness. They conducted a set of tests to investigate the effects of shot peening and polishing on the fatigue strength of case-hardened rollers. Some of the shot-peened rollers were reground and some were polished by a process called barreling. The reground rollers had a roughness average (R_a) of $0.78\text{ }\mu\text{m}$ ($31\text{ }\mu\text{in.}$). The polished rollers had a R_a of $0.05\text{ }\mu\text{m}$ ($2.0\text{ }\mu\text{in.}$). Pitting tests were conducted using a slide-roll ratio of -20 percent on the follower with mineral oil as the lubricant. The lubricant film thickness was estimated to be $0.15 \sim 0.25\text{ }\mu\text{m}$ ($5.9 \sim 9.8\text{ }\mu\text{in.}$). The surface durability of the rollers that

had been shot peened and polished by barreling was significantly improved compared with rollers that were shot peened only or that were shot peened and reground. They found that the pitting limits (maximum Hertz stress with no pitting after 10^7 cycles) of the shot-peened/reground rollers and the shot-peened/polished rollers were 2.15 GPa (312 ksi) and 2.45 GPa (355 ksi), respectively.

Patching, et al. (ref. 4.14) evaluated the scuffing properties of ground and superfinished surfaces using turbine engine oil as the lubricant. The evaluation was performed using case-carburized steel discs. The discs were finish ground in the axial direction such that the orientation of the roughness would be perpendicular to the direction of rolling and sliding, thereby simulating the conditions normally found in gears. Some of the discs were superfinished to provide smoother surfaces. The Ra of the ground discs was about $0.4\text{ }\mu\text{m}$ ($16\text{ }\mu\text{in.}$), and the Ra of the superfinished discs was less than $0.1\text{ }\mu\text{m}$ ($4\text{ }\mu\text{in.}$). They found that compared with the ground discs, the superfinished discs had a significantly higher scuffing load capacity when lubricated with turbine engine oil and subjected to relatively high rolling and sliding speeds. They also noted that under these operating conditions, the sliding friction of the superfinished surfaces was the order of half that for the ground surfaces.

These previous works (refs. 4.1 to 4.14) provide strong evidence that the reduction of surface roughness improves the lubricating condition and offers the possibility of increasing the surface fatigue lives of gears. However, there is little published data to quantify the improvement in life for case-carburized gears. The present study was therefore carried out to quantify the surface fatigue lives of aerospace-quality gears that have been provided with an improved surface finish relative to conventionally ground gears.

4.2 Test Apparatus, Specimens, and Procedure

Gear Test Apparatus.—The gear fatigue tests were performed in the NASA Glenn Research Center's gear test apparatus. The test rig is shown in figure 4.2.1(a) and is described in reference 4.15. The rig uses the four-square principle of applying test loads so that the input drive only needs to overcome the frictional losses in the system. The test rig is belt driven and operated at a fixed speed for the duration of a particular test.

A schematic of the apparatus is shown in figure 4.2.1(b). Oil pressure and leakage replacement flow is supplied to the load vanes through a shaft seal. As the oil pressure is increased on the load vanes located inside one of the slave gears, torque is applied to its shaft. This torque is transmitted through the test gears and back to the slave gears. In this way, power is recirculated and the desired load and corresponding stress level on the test gear teeth may be obtained by adjusting the hydraulic pressure. The two identical test gears may be started under no load, and the load can then be applied gradually. This arrangement also has the advantage that changes in load do not affect the width or position of the running track on the gear teeth. The gears are tested with the faces offset as shown in figure 4.2.1. By utilizing the offset arrangement for both faces of the gear teeth, four surface fatigue tests can be run for each pair of gears.

Separate lubrication systems are provided for the test and slave gears. The two lubrication systems are separated at the gearbox shafts by pressurized labyrinth seals, with nitrogen as the seal gas. The test gear lubricant is filtered through a $5\text{-}\mu\text{m}$ ($200\text{-}\mu\text{in.}$) nominal fiberglass filter. A vibration transducer mounted on the gearbox is used to automatically stop the test rig when gear surface fatigue damage occurs. The gearbox is also automatically stopped if there is a loss of oil flow to either the slave gearbox or the test gears, if the test gear oil overheats, or if there is a loss of seal gas pressurization.

Test Specimens.—The gears of the present study were manufactured from consumable-electrode vacuum-melted (CVM) AISI 9310 steel. The best available baseline for this study is a set of conventionally ground gears that were previously tested and the data reported (ref. 4.6). The test gears of the present study and those of the baseline study were manufactured from separate melts of consumable-electrode vacuum-melted (CVM) AISI 9310 steel. Both sets of gears were case carburized and ground. The nominal and certified chemical compositions of the gears are given in table 4.2.1. Figure 4.2.2(a) to (d) are photomicrographs showing the microstructure of the case and core. Figure 4.2.3 is a plot of

material hardness versus depth below the pitch radius surface. The data of figure 4.2.3 are equivalent Rockwell C scale hardness values converted from Knoop microhardness data. The differences of the curves displayed on figure 4.2.3 are typical of sample-to-sample differences for case-carburized gears. These data and metrology inspections (ref. 4.16) verify that the gear materials and geometry are aerospace quality.

The dimensions of the gears are given in table 4.2.2. The gears are 3.175 mm module (8 diametral pitch) and have a standard 20° involute pressure angle with tip relief of 0.013 mm (0.0005 in.) starting at the highest point of single tooth contact. The nominal face width is 6.35 mm (0.250 in.), and the gears have a nominal 0.13-mm (0.005-in.) radius edge break to avoid edge loading.

Fourteen gears were selected for finishing by a polishing method described below. A subset of four gears was selected at random for metrology inspections, both before and after superfinishing. Parameters measured on each gear were lead and profile errors, adjacent pitch errors, and mean circular tooth thickness. In order to show the detailed effects of superfinishing, it was decided to also take “relocated” profiles from the gear teeth. This was achieved by use of a special stepper-motor-driven profilometer with which it was possible to take a profile or series of profiles at a precisely known location on a gear tooth. The principle of relocation was based on detection of the edges of the tooth by running the profilometer stylus in the axial direction of the gear to detect the side of the tooth and radially to detect the tooth tip. Three profiles were taken from both sides of two teeth on each gear (i.e., a total of 12 profiles from each gear). Two of the three profiles on each gear flank were located one mm (0.039 in.) from each side edge and the third profile was located on the center of the tooth. Profile data was taken up to and slightly beyond the tip of the teeth as a direct means of verifying the accuracy of relocation in every case. All profiles were processed using a standard phase-corrected digital filter with a cutoff of 0.08 mm (0.003 in.).

Superfinishing treatment of the gears was completed as follows. The gears were immersed in a bed of small zinc chips, water, and aluminum oxide powder. The container (a rubber-lined open tank) was vibrated for a period of several hours and the grade of the oxide powder was increased in fineness in three stages. Upon completion of the initial superfinish treatment, metrology inspections were carried out and relocated profiles were taken. Although the surface finish had been improved, grinding marks were still visible on some teeth. The gears were then subjected to a second superfinish treatment. After the second treatment, the gears had a superb near-mirror finish (fig. 4.2.4), and grinding marks were no longer visible. Following the second (final) superfinish treatment, metrology and profilometry inspections were again completed. A detailed report of the superfinish treatment and inspections is available (ref. 4.16). From analysis of the metrology data, it was concluded that the superfinishing treatment did not significantly alter the lead and involute profile traces of the gear teeth.

Figure 4.2.5 is a typical comparison of the relocated surface profiles of the same tooth taken first after grinding, a second time after the initial superfinish treatment, and a third time after the final superfinish treatment. The profile taken after the first stage of superfinishing (fig. 4.2.5(b)) shows a persistence of identifiable grinding marks. These have almost disappeared from the profile taken after the final superfinish treatment (fig. 4.2.5(c)), although there are faint signs of particularly deep marks. Analysis of the profilometry data suggested that about 1 μm (39 $\mu\text{in.}$) had been removed from each surface following the initial superfinish treatment and in total, about 2 to 3 μm (79 to 118 $\mu\text{in.}$) had been removed from the surface following the final stage of treatment. These estimates of material removed, as derived from the profilometry data, agree with estimates obtained from metrology measurements of the mean circular tooth thickness taken before and after finishing (ref. 4.16). The roughness average (R_a) and 10-point parameter (R_z) values for each profile inspection were calculated using the profilometry data filtered with a cutoff of 0.08 mm (0.003 in.). Table 4.2.3 is a statistical summary of the calculated R_a and R_z values. Before superfinishing, the gears had a mean R_a of 0.380 μm (15 $\mu\text{in.}$) and a mean R_z of 3.506 μm (138 $\mu\text{in.}$). After superfinishing, the gears had a mean R_a of 0.071 μm (2.8 $\mu\text{in.}$) and a mean R_z of 0.940 μm (37 $\mu\text{in.}$). Therefore, the mean R_a and mean R_z values were reduced by a factor of about 5 and 4, respectively, by superfinishing.

A ground gear tooth and a superfinished gear tooth were inspected using a mapping interferometric microscope. Data from the microscope were low pass filtered to remove instrument noise and were further processed to remove the datum. Figure 4.2.6 is a comparison of the processed interferometric data. The images of figure 4.2.6(a) and (b) are not images of the same gear before and after superfinishing but are images from two separate gears. These images provide examples of features of typical ground and superfinished surfaces. Figure 4.2.6(b) shows that traces of the original grinding marks are still evident after superfinishing, but the depths of the marks are greatly reduced.

Test Procedure.—The lubricant used was developed for helicopter gearboxes under the specification DOD-L-85734. This is a 5-cSt lubricant of a synthetic polyol-ester base stock with an antiwear additive package. Lubricant properties gathered from references 4.7 and 4.17 are provided in table 4.2.4.

The test gears were run with the tooth faces offset by a nominal 3.3 mm (0.130 in.) to give a surface load width on the gear face of 3.0 mm (0.120 in.). The actual tooth face offset for each test is based on the measured face width of the test specimen, and the offset is verified upon installation using a depth gage. The nominal 0.13-mm (0.005-in.) radius edge break is allowed for to calculate load intensity. All tests were run-in at a load (normal to the pitch circle) per unit width of 123 N/mm (700 lb/in.) for 1 hour. The load was then increased to 580 N/mm (3300 lb/in.), which resulted in a 1.71-GPa (248-ksi) pitch-line maximum Hertz stress. At the pitch-line load, the tooth bending stress was 0.21 GPa (30 ksi) if plain bending was assumed. However, because there was an offset load, there was an additional stress imposed on the tooth bending stress. The combined effects of the bending and torsional moments yield a maximum stress of 0.26 GPa (37 ksi). The effects of tip relief and dynamic load were not considered for the calculation of stresses.

The gears were tested at 10,000 rpm, which gave a pitch-line velocity of 46.5 m/s (9154 ft/min). Inlet and outlet oil temperatures were continuously monitored. Lubricant was supplied to the inlet of the gear mesh at 0.8 liter/min (49 in.³/min) and 320±7 K (116±13 °F). The lubricant outlet temperature was recorded and observed to have been maintained at 348±4.5 K (166±8 °F). The tests ran continuously (24 hr/day) until a vibration detection transducer automatically stopped the rig. The transducer is located on the gearbox adjacent to the test gears. If the gears operated for 500 hours (corresponding to 300 million stress cycles) without failure, the test was suspended. The lubricant was circulated through a 5-μm (200-μin.) nominal fiberglass filter to remove wear particles. For each test, 3.8 liter (1 gal) of lubricant was used.

The film thickness at the pitch point for the operating conditions of the surface fatigue testing was calculated using the computer program EXTERN. This program, developed at the NASA Glenn Research Center, is based on the methods of References 4.18 and 4.19. For the purposes of the calculation, the gear surface temperature was assumed to be equal to the average oil outlet temperature. This gave a calculated pitch-line film thickness of 0.54 μm (21 μin.)

4.3 Results and Discussion

Surface fatigue testing was completed on a set of gears manufactured from CVM AISI 9310 steel. The gears were case carburized, ground, and superfinished. The measured Ra of the superfinished gears was 0.071 μm (2.8 μin.). Gear pairs were tested until failure or until 300 million stress cycles (500 hr of testing) had been completed with no failure. The test conditions were a load per unit width of 580 N/mm (3300 lb/in.), which resulted in a 1.71-GPa (248-ksi) pitch-line maximum Hertz stress. For purposes of this work failure was defined as one or more spalls or pits covering at least 50 percent of the width of the Hertzian line contact on any one tooth. Examples of fatigue damage are shown in figure 4.3.1. Figure 4.3.1 also provides scaled measures of the running tracks. The actual widths of the running tracks varied slightly from test to test depending on the exact geometry of the edge break radius provided to prevent edge loading. Damage occurred after longer running times for the superfinished gears relative to ground gears. The damaged gear surfaces had similar features regardless if the gear was superfinished or ground.

To provide a baseline for the present study, the data from reference 4.7 were selected as the most appropriate available. The tests of reference 4.7 were conducted using the same rigs, lubricant,

temperatures, loads, speeds, material specification, and geometry specifications as the present study. The gears of the baseline study were specified to be ground with a maximum root-mean-squared roughness of $0.406\text{-}\mu\text{m}$ ($0.016\text{-}\mu\text{in}$). There were 17 failures and 3 suspended tests for the ground gears of the baseline study, and there were 8 failures and 7 suspended tests for the ground and superfinished gears of the present study. The test data were analyzed by considering the life of each pair of gears as a system. The data were analyzed using the maximum likelihood method described in chapter 3.

Surface fatigue test results for the ground gears of the baseline study are shown in figure 4.3.2(a). The line shown on figure 4.3.2(a) is a maximum likelihood fit of the data to a two-parameter Weibull distribution. From the fit line, the 10- and 50-percent lives of the sample population are 7.6×10^6 and 67×10^6 stress cycles. Surfaces that had been run but were not pitted or spalled had a different appearance relative to the appearance before testing. The grinding marks had become worn away and/or smeared, and the running tracks on the gears were plainly evident (fig. 4.3.1).

Surface fatigue test results for the ground and superfinished gears of the present study are shown in figure 4.3.2(b). The line shown on figure 4.3.2(b) is a maximum likelihood linear fit of the data to a two-parameter Weibull distribution. From the fit line, the 10- and 50-percent lives of the sample population are 43×10^6 and 265×10^6 stress cycles. Superfinished surfaces that had been run and survived with no fatigue failure appeared almost like surfaces that had not been run. The running tracks on the gears were not immediately evident, but slight traces could be seen by close examination with a 10X eyepiece. The wear and/or smearing that were seen on the ground gears after testing were not observed on the tested superfinished gears.

The surface fatigue test results are summarized in table 4.3.1 and figures 4.3.2(c) and (d). Figure 4.3.2(c) shows the two maximum likelihood fit lines on one Weibull plot. The Weibull slopes are similar, and therefore the gears have similar relative failure distributions. Figure 4.3.2(d) shows the distributions of fatigue lives plotted using linear axes. This plot shows that for a given reliability, the lives of the superfinished gears are greater than the lives of the ground gears. One significant result of the statistical analysis is that the 10-percent life of the set of ground and superfinished gears was greater than the 10-percent life of the set of ground gears to a 90-percent confidence level. In general, the life of the set of ground and superfinished gears was about five times greater than the life of the set of ground gears. In this study, the difference in life can be attributed to the combined effects of (a) the gears being made from different melts of steel and (b) the superfinished gear teeth surface having significantly different topographies.

To help assess the influence of the superfinishing on life, the results of the present study can be compared in a qualitative sense to the NASA Glenn gear fatigue database. The data were gathered from several publications (refs. 4.7, 4.20 to 4.27). The data in these reports had been reported graphically. The data plots were digitized to determine the numeric values, and the data were analyzed using the maximum likelihood methods described in chapter 3. The data are summarized in table 4.3.2. These data represent the majority of testing of AISI 9310 gears using the NASA Glenn gear fatigue test apparatus (fig. 4.2.1). Common to all data presented in table 4.3.2 are:

1. tests completed using the same set of four rigs;
2. test gear geometry per Table 4.2.2;
3. ground surface finish specified as maximum root-mean-squared roughness of $0.406\text{-}\mu\text{m}$ ($0.016\text{-}\mu\text{in}$);
4. load of 1.71-GPa (248-ksi) Hertz contact stress at the pitch line;
5. test gears run in an offset condition with a 3.3-mm (0.130-in.) nominal tooth surface overlap;
6. operating speed of 10 000 rpm;
7. lubricant filtered using a $5\text{-}\mu\text{m}$ ($200\text{-}\mu\text{in.}$) nominal filter to remove wear debris
8. lubricant outlet temperature maintained at $348 \pm 4.5\text{ K}$ ($166 \pm 8\text{ }^\circ\text{F}$);
9. the test data treated as failures of a system of two gears and then fitted to a two-parameter Weibull distribution using the maximum likelihood method.

The 10- and 50-percent lives listed in table 4.3.2 are from the maximum likelihood fits to the distribution. The table is sorted in ascending order of 10-percent lives, except the data of the present study occupy the last row of the table. The data of table 4.3.2 were produced using gears manufactured from several melts of steel, having various processing (such as shot peening), and lubricated with several different lubricants with viscosities (at 373 K (212 °F)) ranging from 5.1~7.7 cSt. The ground and superfinished gears of the present study had lives greater than those of any other set of single-vacuum processed AISI 9310 gears tested to date. The lives of the CVM AISI 9310 superfinished gears were of the order of magnitude of ground VIM-VAR (vacuum-induction-melted vacuum-arc-remelted) AISI 9310 gears. The increase in fatigue life resulting from the use of VIM-VAR melting practice rather than CVM melting practice is well established. It appears that superfinishing offers a performance improvement, relative to ground surfaces, of similar magnitude to the improvement that VIM-VAR processing offers over CVM processing. Lastly, the proportion of the gears operating for 300 million cycles without failure was considerably higher than that for any of the other gears tested.

Considering the quantitative differences in the data of table 4.3.1, the qualitative comparisons made using the data of table 4.3.2, and the observed differences in appearances of the tested ground and superfinished surfaces, there is strong evidence that superfinishing significantly improves the surface fatigue lives of case-carburized and ground aerospace-quality AISI 9310 gears.

4.4 Conclusions

A set of consumable-electrode vacuum-melted (CVM) AISI 9310 steel gears were ground and then provided with a near-mirror quality tooth surface by superfinishing. The gear teeth surface qualities were evaluated using metrology inspections, profilometry, and a mapping interferometric microscope. The gears were tested for surface fatigue in the NASA Glenn gear fatigue test apparatus at a load of 1.71 GPa (248 ksi) and at an operating speed of 10 000 rpm until failure or until survival of 300 million stress cycles. The lubricant used was a polyol-ester base stock meeting the specification DOD-L-85734. The failures were considered as failures of a two-gear system, and the data were fitted to a two-parameter Weibull distribution using the maximum likelihood method. The results of the present study were compared with the NASA Glenn gear fatigue data base. The following results were obtained.

1. The superfinishing treatment removed about 2 to 3 μm (79 to 118 $\mu\text{in.}$) of material from the tooth surfaces.
2. The superfinishing treatment reduced the mean roughness average (Ra) by a factor of about five and the mean 10-point parameter (Rz) value by a factor of about four.
3. The 10-percent life of the set of ground and superfinished gears of the present study was greater than the 10-percent life of the set of ground gears of the baseline study to a 90-percent confidence level.
4. In general, the life of the set of ground and superfinished gears of the present study was about five times greater than the life of the set of ground gears of the baseline study.
5. The set of ground and superfinished gears of the present study had lives greater than those of any other set of single-vacuum processed AISI 9310 gears tested to date using the NASA Glenn gear fatigue test apparatus.
6. The proportion of the gears operating for 300 million cycles without failure was considerably higher for the superfinished gears than was the proportion for any other set of ground AISI 9310 gears tested to date using the NASA Glenn gear fatigue test apparatus.
7. The lives of the CVM AISI 9310 ground and superfinished gears of the present study were of the order of magnitude of VIM-VAR AISI 9310 ground gears when tested using the NASA Glenn gear fatigue test apparatus.
8. There is strong evidence that superfinishing significantly improves the surface fatigue lives of case-carburized, ground, aerospace-quality AISI 9310 gears.

References

- 4.1. Anderson, W.; Parker, R.; Zaretsky, E., "Effect of Nine Lubricants on Rolling-Contact Fatigue Life," *NASA TN D-1404*, (1962).
- 4.2. Anderson, W.; Sibley, L.; Zaretsky, E., "The Role of Elastohydrodynamic Lubrication in Rolling-Contact Fatigue," *J. of Basic Eng.*, **85**, (1963).
- 4.3. Liu, J.; Tallian, T.; McCool, J., "Dependence of Bearing Fatigue Life on Film Thickness to Surface Roughness Ratio," *ASLE Trans.*, **18 [2]**, (1975).
- 4.4. Skurka, J., "Elastohydrodynamic Lubrication of Roller Bearings," *Trans. Of the ASME, Journal of Lubrication Technology*, (1970).
- 4.5. Borsoff, V., "On the Mechanism of Gear Lubrication," *J. of Basic Eng.*, (1959).
- 4.6. Townsend, D.; Shimski, J., "Evaluation of Advanced Lubricants for Aircraft Applications Using Gear Surface Fatigue Tests," *NASA TM-104336*, (1991).
- 4.7. Townsend, D.; Shimski, J., "Evaluation of the EHL Film Thickness and Extreme Pressure Additives on Gear Surface Fatigue Life," *NASA TM-106663*, (1994).
- 4.8. Tanaka, S.; Ishibashi, A.; Ezoe, S., "Appreciable Increases in Surface Durability of Gear Pairs With Mirror-Like Finish," *ASME Pub. 84-DET-223*, ASME, New York, NY, (1984).
- 4.9. Ishibashi, A.; Ezoe, S.; Tanaka, S., "Mirror Finishing of Tooth Surfaces Using a Trial Gear Grinder With Cubic-Born-Nitride Wheel," *ASME Pub. 84-DET-153*, ASME, New York, NY, 1984.
- 4.10. Nakatsuji, T.; Mori, A.; Shimotsuma, Y., "Pitting Durability of Electrolytically Polished Medium Carbon Steel Gears," *Tribology Trans.*, **38 [2]**, (1995) 223–232.
- 4.11. Nakatsuji, T.; Mori, A., "Pitting Durability of Electrolytically Polished Medium Carbon Steel Gears—Succeeding Report," *Tribology Trans.*, **42 [2]**, (1999) 393–400.
- 4.12. Hoyashita, S.; Hashimoto, M.; Seto, K., "Basic Studies on Fatigue Strength of Case-Hardened Gear Steel—Effects of Shot Peening and/or Barreling Processes," *AGMA Pub. 98FTM3*, (1998).
- 4.13. Hashimoto, M.; Hoyashita, S.; Iwata, J., "Studies on Improvement of Surface Durability of Case-Carburized Steel Gear—Effects of Surface Finish Processes upon Oil Film Formation," *AGMA Pub. 98FTM9*, (1998).
- 4.14. Patching, M.; Kweh, C.; Evans, H.; Snidle, R., "Conditions for Scuffing Failure of Ground and Superfinished Steel Disks at High Sliding Speeds Using a Gas Turbine Engine Oil," *J. of Tribology*, **117 [3]**, (1995) 482–489.
- 4.15. Townsend, D.; Chevalier, J.; Zaretsky, E., "Pitting Fatigue Characteristics of AISI M-50 and Super Nitralloy Spur Gears," *NASA TN D-7261*, (1973).
- 4.16. Snidle, R.; Evans, H.; Alanou, M., "The Effect of Superfinishing on Gear Tooth Profile," *AD-A327916*, Defense Technical Information Center, Ft. Belvoir, VA, (1997). Also available from NASA Scientific and Technical Information Center, <http://www.sti.nasa.gov/> accessed October 3, 2000.
- 4.17. Present, D.; Newman, F.; Tyler, J.; Cuellar, J., "Advanced Chemical Characterization and Physical Properties of Eleven Lubricants," *NASA CR-168187*, (1983).
- 4.18. Anderson, N.; Lowenthal, S.; and Black, J., "An Analytical Method to Predict Efficiency of Aircraft Gearboxes," *NASA TM-83716, AIAA Paper 84-1500*, (1984).
- 4.19. Anderson, N.; Lowenthal, S., "Spur-Gear-System at Part and Full Load," *NASA TP-1622*, (1980).
- 4.20. Townsend, D.; Zaretsky, E., "Effect of Five Lubricants on Life of AISI 9310 Spur Gears," *NASA TP-2419*, (1985).
- 4.21. Coy, J.; Townsend, D.; Zaretsky, E., "Analysis of Dynamic Capacity of Low-Contact-Ratio Spur Gears Using Lundberg-Palmgren Theory," *NASA TN D-8029*, (1975).
- 4.22. Scibbe, W.; Townsend, D.; Aron, P., "Effect of Lubricant Extreme-Pressure Additives on Surface Fatigue Life of AISI 9310 Spur Gears," *NASA TP-2408*, (1984).
- 4.23. Townsend, D., "Surface Fatigue Life of High Temperature Gear Materials," *NASA TM-106562, AIAA Paper 94-3078*, (1994).

- 4.24. Townsend, D.; Parker, R.; Zaretsky, E., "Evaluation of CBS 600 Carburized Steel as a Gear Material," *NASA TP-1390*, (1979).
- 4.25. Townsend, D.; Zaretsky, E., "Effect of Shot Peening on Surface Fatigue Life of Carburized and Hardened AISI 9310 Spur Gears," *NASA TP-2047*, (1982).
- 4.26. Townsend, D., "Improvements in Surface Fatigue Life of Hardened Gears by High-Intensity Shot Peening," *NASA TM-105678*, (1992).
- 4.27. Townsend, D.; Bamberger, E., "Surface Fatigue Life of Carburized and Hardened M50NiL and AISI 9310 Spur Gears and Rolling-Contact Test Bars," *NASA TM-101979*, (1989).

TABLE 4.2.1.—NOMINAL AND CERTIFIED CHEMICAL COMPOSITION OF GEAR MATERIALS, AISI 9310

	Element									
	C	Mn	P	S	Si	Ni	Mo	Cr	Cu	Fe
Nominal contents, wt%	0.10	0.63	0.005	0.005	0.27	3.22	0.12	1.21	0.13	Balance
Ground gear, certified contents, wt%	0.10	0.56	0.003	0.003	0.26	3.49	0.10	1.15	*	*
Superfinished gear, certified contents, wt%	0.11	0.55	0.006	0.018	0.26	3.42	0.10	1.30	*	*

*Indicates not measured.

TABLE 4.2.2.—SPUR GEAR DATA.
[GEAR TOLERANCE PER AGMA CLASS 12]

Number of teeth	28
Module, mm	3.175
Diametral pitch	8
Circular pitch, mm (in.)	9.975 (0.3927)
Whole depth, mm (in.)	7.62 (0.300)
Addendum, mm (in.)	3.18 (.125)
Chordal tooth thickness reference, mm (in.)	4.85 (0.191)
Pressure angle, deg.	20
Pitch diameter, mm (in.)	88.90 (3.500)
Outside diameter, mm (in.)	95.25 (3.750)
Root fillet, mm (in.)	1.02 to 1.52 (0.04 to 0.06)
Measurement over pins, mm (in.)	96.03 to 96.30 (3.7807 to 3.7915)
Pin diameter, mm (in.)	5.49 (0.216)
Backlash reference, mm (in.)	0.254 (0.010)
Tip relief, mm (in.)	0.010 to 0.015 (0.0004 to 0.0006)

TABLE 4.2.3.—SUMMARY OF STATISTICAL ANALYSIS OF PROFILOMETRY DATA

Parameter	Surface condition	Mean value, μm (μin.)	Standard deviation, μm (μin.)
Roughness average (Ra)	Before superfinishing	0.380 (15.0)	0.068 (2.7)
	After superfinished	0.070 (2.8)	0.016 (0.6)
10-point parameter (Rz)	Before superfinishing	3.506 (138.0)	0.610 (24.0)
	After superfinished	0.940 (37.0)	0.298 (11.7)

^aData are based on relocated and filtered profile measurements of the same teeth, both before and after superfinishing.TABLE 4.2.4.—LUBRICANT PROPERTIES.
(FROM REFS. 4.7 AND 4.17)

Specification	DOD-L-85734
Basestock	Polyol-ester
Kinematic viscosity, cSt	
311 K (100 °F)	27.6
372 K (210 °F)	5.18
Absolute viscosity, N.s/m ²	
333 K (140 °F)	0.01703
355 K (180 °F)	0.00738
372 K (210 °F)	0.00494
Specific gravity	
289 K (60 °F)	0.995
372 K (210 °F)	0.954
Pressure viscosity coefficient (1/Pa)	
313 K (104 °F)	11.4 x 10 ⁻⁹
373 K (212 °F)	9.5 x 10 ⁻⁹
Total acid number (tan), Mg Koh/g oil	0.40
Flash point, K (°F)	544 (520)
Pour point, K (°F)	211 (-80)

TABLE 4.3.1.—FATIGUE LIFE RESULTS FOR TEST GEARS

Gears	10-percent life, cycles	50-percent life, cycles	Weibull slope	Failure index ^a	Confidence number, ^b percent
CVM AISI 9310, ground (ref. 4.7)	7.6×10^6	67×10^6	0.9	17/20	---
CVM AISI 9310, superfinished	43×10^6	265×10^6	1.0	8/15	91

^a Indicates the number of failures out of the number of tests.

^b Probability, expressed as a percentage, that the 10-percent life of the superfinished gears is greater than the 10-percent life of the ground gears.

TABLE 4.3.2.—SURFACE FATIGUE LIVES OF CASE-CARBURIZED AISI 9310 GEAR PAIRS TESTED
IN THE NASA GLENN RESEARCH CENTER GEAR FATIGUE TEST APPARATUS

[Pitch-line Hertz stress, 1.71 GPa (248 ksi), 10 000 rpm; oil outlet temperature maintained at 348±4.5 K (166±8 °F); 5-µm (200-µin.) nominal fiberglass filter to remove wear debris; gear geometry, AGMA class 12, 8 pitch, 28 teeth, AISI 9310 gear steel.]

Ref.	Year published	Melting method	Lubricant specification or description	10-percent life, cycles ^a	50-percent life, cycles ^a	Weibull slope	Failure index ^b	Comments
4.20	1985	CVM	MIL-L-23699	4.3×10 ⁶	26×10 ⁶	1.0	20/20	Ground
4.21	1975	CVM	Super-refined mineral oil & additives	6.8×10 ⁶	14×10 ⁶	2.6	19/19	Ground
4.7	1994	CVM	DOD-L-85734	7.6×10 ⁶	67×10 ⁶	0.9	17/20	Ground, most direct baseline for the present study
4.22	1984	CVM	Tetraester plus additives (see ref.)	9.9×10 ⁶	74×10 ⁶	0.9	22/23	Ground
4.22	1984	CVM	Tetraester plus additives (see ref.)	11×10 ⁶	32×10 ⁶	1.7	30/30	Ground
4.7	1994	CVM	MIL-L-23699	12×10 ⁶	71×10 ⁶	1.0	20/20	Ground
4.23	1994	CVM	NASA standard	18×10 ⁶	44×10 ⁶	2.2	18/19	Ground
4.22	1984	CVM	Tetraester plus additives (see ref.)	18×10 ⁶	75×10 ⁶	1.3	21/21	Ground
4.20	1985	CVM	Dibasic acid ester gear lubricant	19×10 ⁶	41×10 ⁶	2.4	20/20	Ground
4.22	1984	CVM	Tetraester plus additives (see ref.)	19×10 ⁶	65×10 ⁶	1.5	20/20	Ground
4.24	1979	CVM	NASA standard	22×10 ⁶	54×10 ⁶	2.1	30/30	Ground
4.20	1985	CVM	MIL-L-23699 – type II	25×10 ⁶	39×10 ⁶	4.2	20/20	Ground
4.20	1985	CVM	GM 6137-M	25×10 ⁶	53×10 ⁶	2.6	20/20	Ground
4.25	1982	CVM	NASA standard	26×10 ⁶	66×10 ⁶	2.0	24/24	Ground, shot peened.
4.26	1992	VIM-VAR	NASA standard	32×10 ⁶	170×10 ⁶	1.1	14/20	Ground, medium-intensity shot peened
4.27	1989	VIM-VAR	NASA standard	52×10 ⁶	197×10 ⁶	1.4	24/33	Ground
4.26	1992	VIM-VAR	NASA standard	89×10 ⁶	250×10 ⁶	1.8	13/20	Ground, high-intensity shot peened
N/A	2000	CVM	DOD-L-85734	43×10 ⁶	265×10 ⁶	1.0	8/15	Ground and superfinished (present study)

^aThe 10-percent and 50-percent lives are those obtained by fitting the test data to two-parameter Weibull distributions, the lives are system lives, the system being a pair of gears.

^bIndicates the number of failures out of the number of tests. A test was suspended after 300×10⁶ cycles if no failure occurred

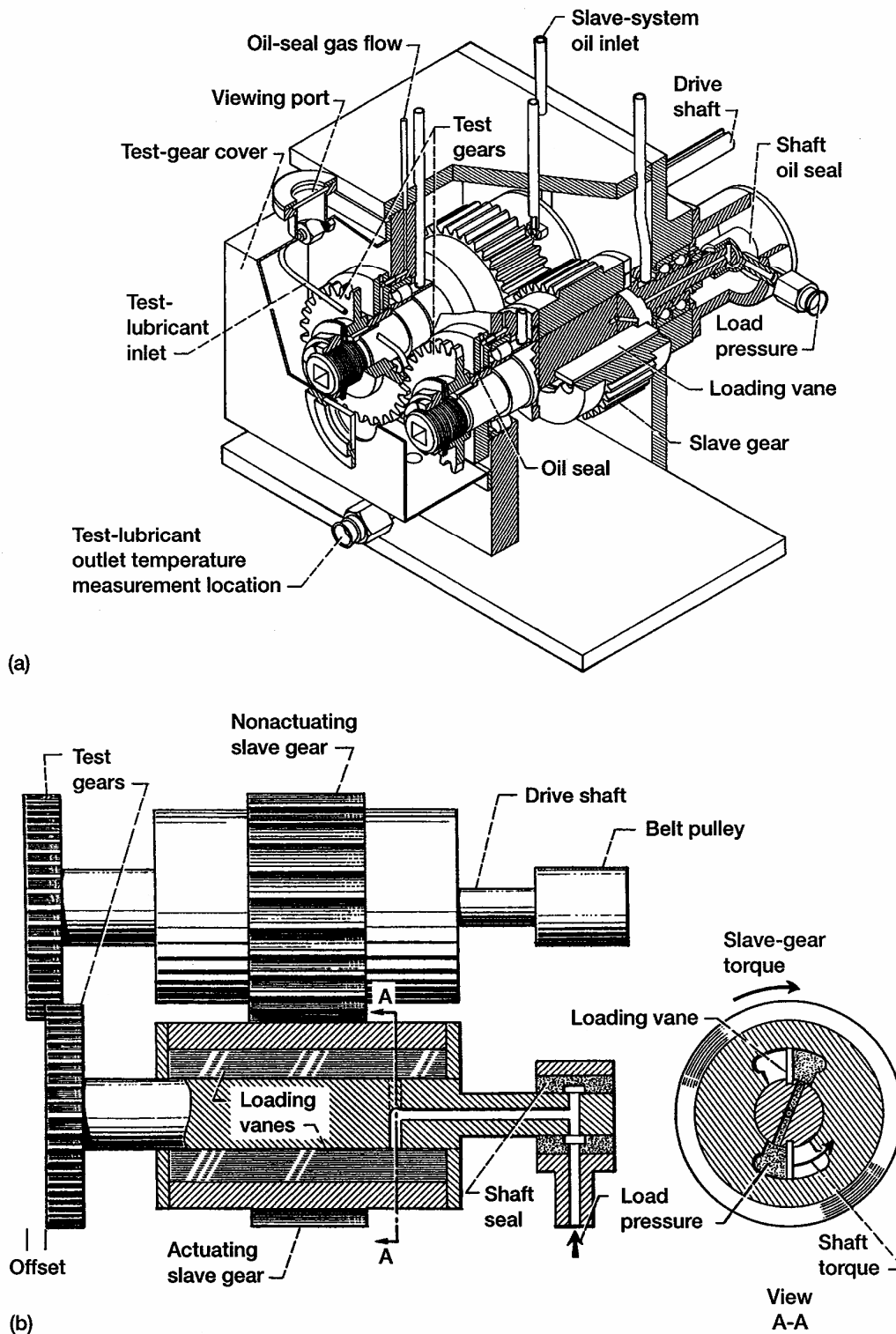
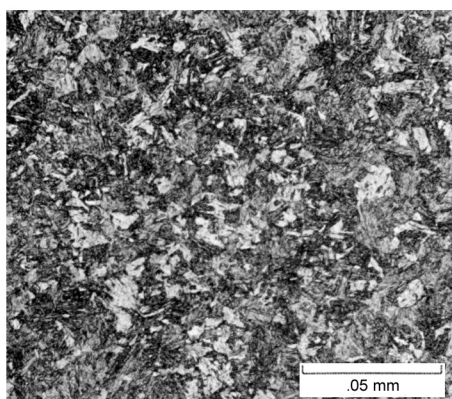
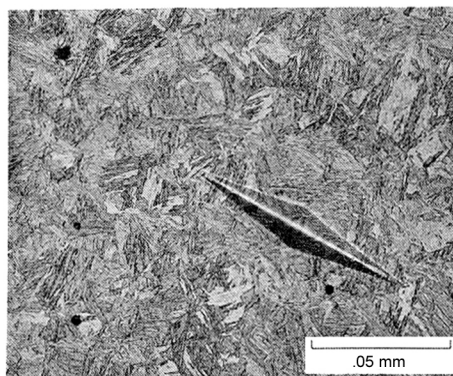


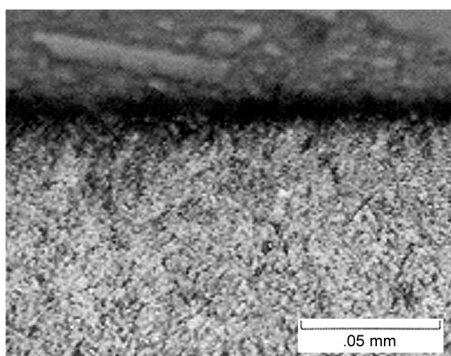
Figure 4.2.1.—NASA Glenn Research Center gear fatigue test apparatus. (a) Cutaway view.
(b) Schematic view.



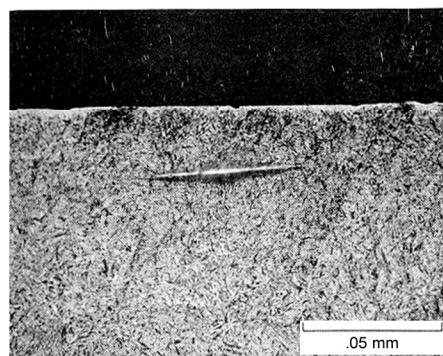
(a)



(c)



(b)



(d)

Figure 4.2.2.—Microphotographs of the gears prepared with 3 percent nital etch. (a) Core of superfinished gear. (b) Case of superfinished gear. (c) Core of ground gear (from ref. 4.7) (d) Case of ground gear (from ref. 4.7).

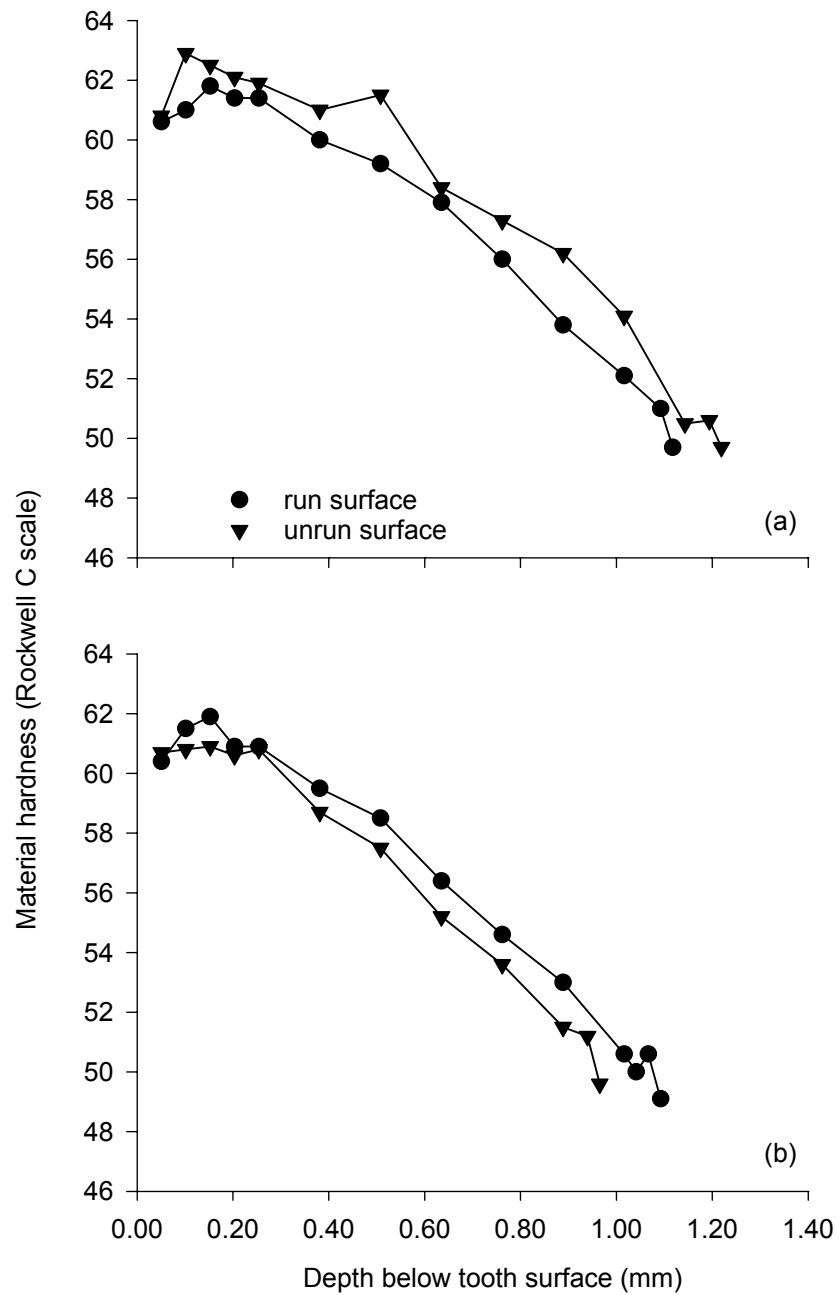


Figure 4.2.3.—Material hardness versus depth below the pitch radius surface. (a) Superfinished gear; (b) Ground gear.



Figure 4.2.4.—Near-mirror quality of superfinished tooth surface.

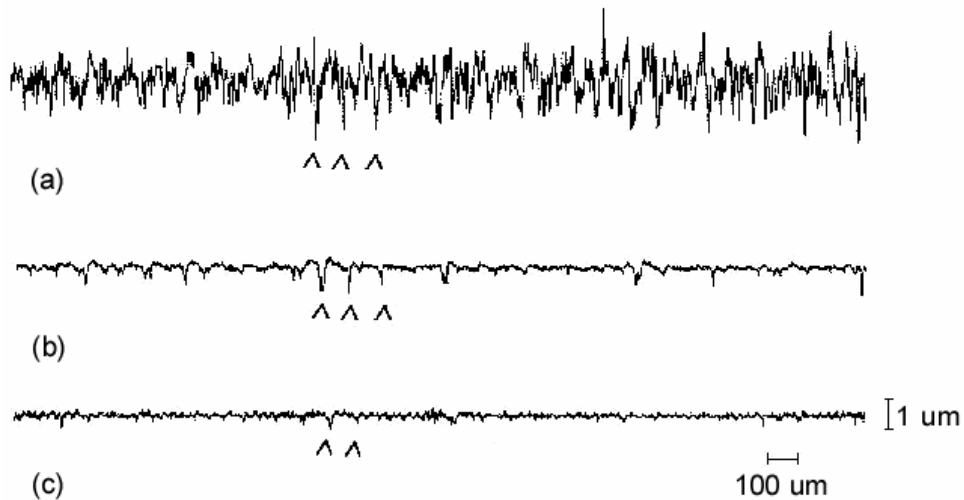


Figure 4.2.5.—Typical relocated surface features measured using a profilometer followed by filtering of the data using a 0.08-mm (0.003-in.) cutoff. Evidence of persistence of the deepest grinding marks are indicated by arrows.
 (a) Ground tooth surface, $R_a = 0.434 \mu\text{m}$ (17 $\mu\text{in.}$). (b) Same tooth surface after the first stage of superfinishing, $R_a = 0.083 \mu\text{m}$ (3.3 $\mu\text{in.}$). (c) Same tooth after second (final) stage of superfinishing, $R_a = 0.056 \mu\text{m}$ (2.2 $\mu\text{in.}$).

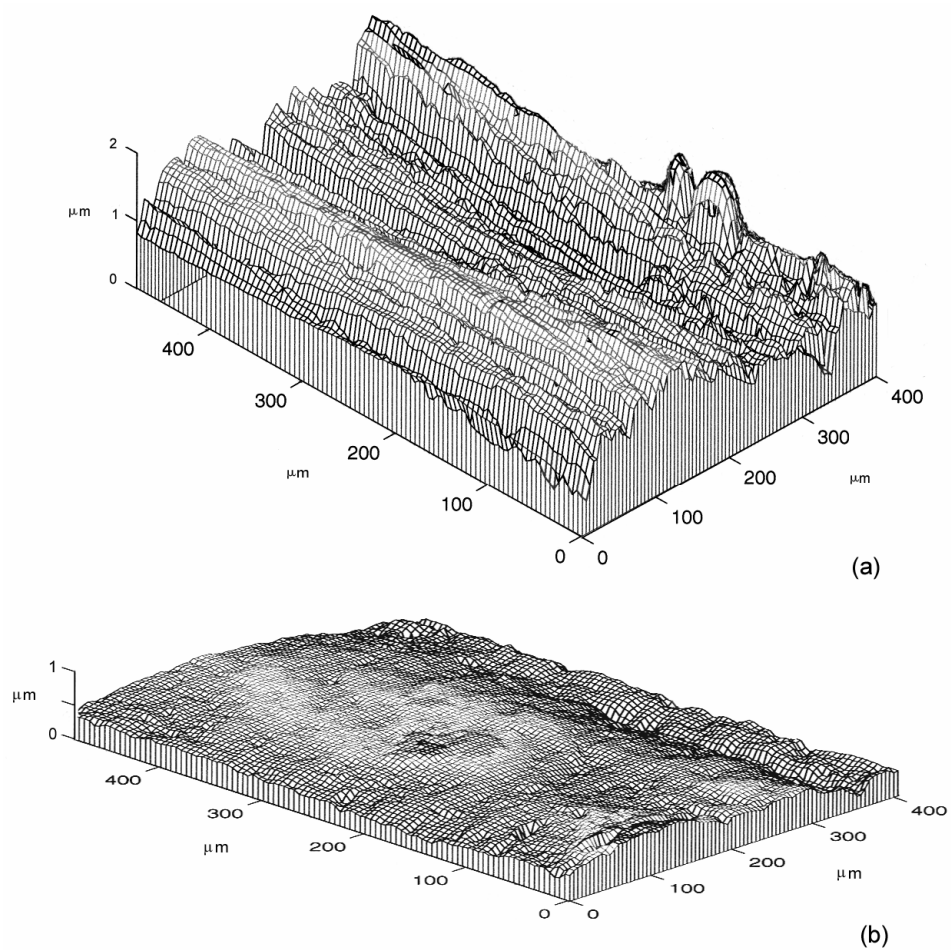


Figure 4.2.6.—Comparison of gear tooth surface topographies as measured using a mapping interferometric microscope. (a) Ground gear tooth. b) Superfinished gear tooth.

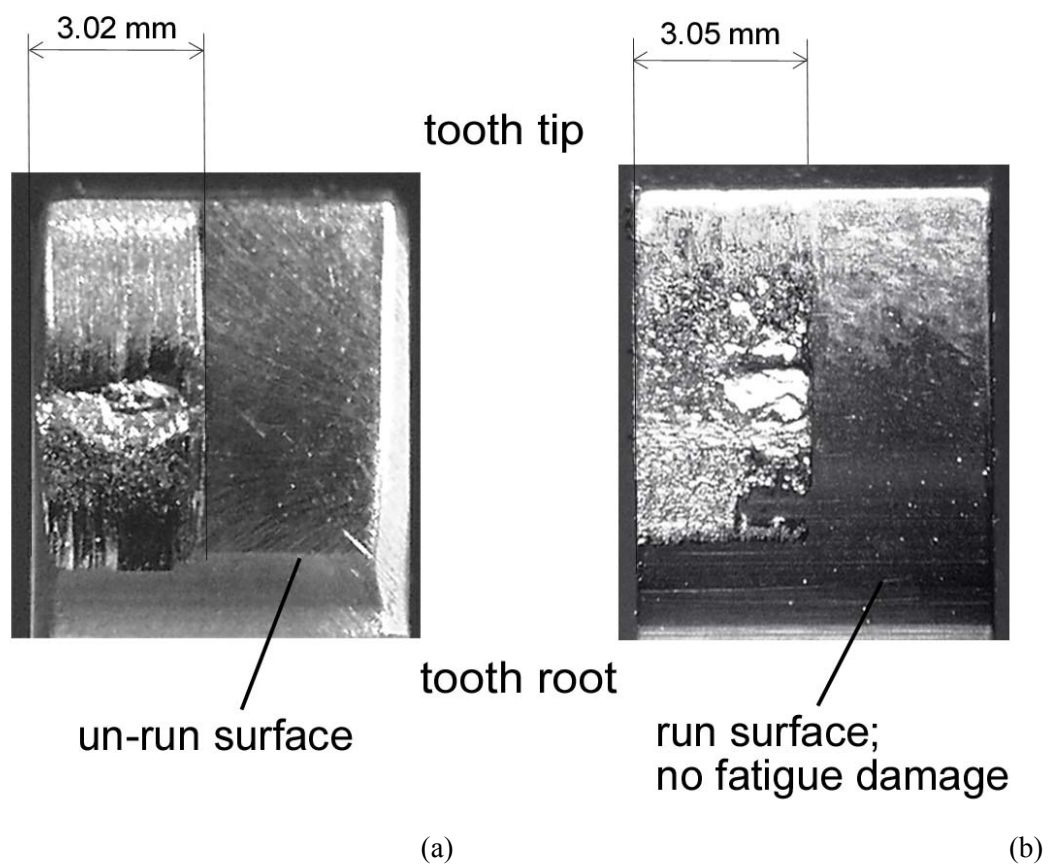


Figure 4.3.1.—Typical fatigue damage. (a) Ground gear from study of reference 4.7.
(b) Superfinished gear of present study.

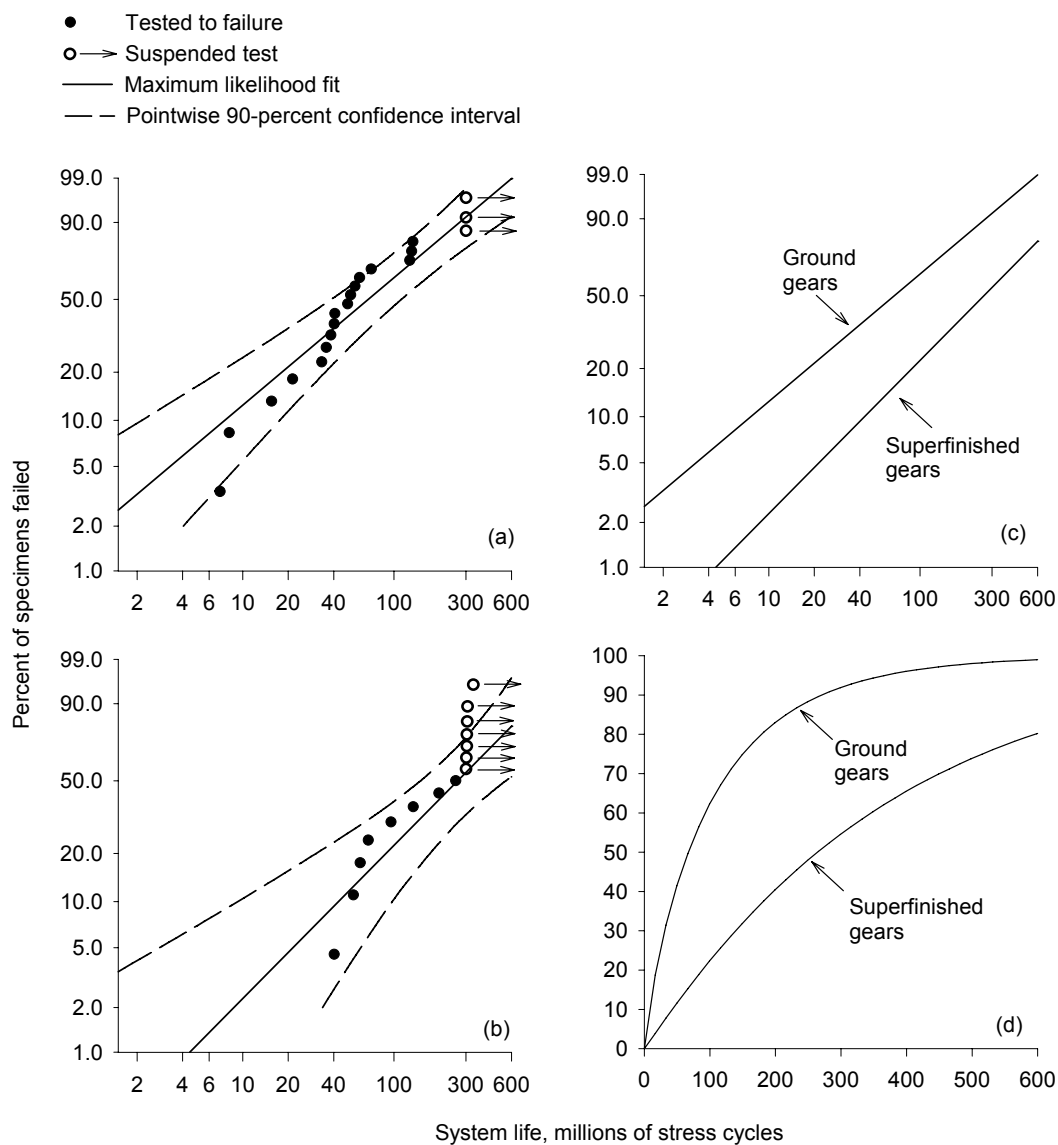


Figure 4.3.2.—Surface fatigue lives of ground and superfinished AISI 9310 gear pairs. (a) Ground gears. (b) Superfinished gears. (c) Summary of maximum likelihood fit lines. (d) Maximum likelihood fit lines plotted on linear axes. Individual data points are displayed using exact median rank plotting positions.

Chapter 5—Evaluation of the Experimental Conditions (Methodology)

5.1 Introduction

The experiments documented in chapter 4 showed conclusively that gears with differing as-manufactured surface topographies can have dramatically differing performance characteristics. To apply these laboratory evaluations to products requires engineering understanding, analysis and judgment. In this chapter methodologies for evaluating the experimental condition are developed so as to assist with the application of superfinishing for gears.

It is well known that the dynamic gear tooth forces can deviate significantly from the static equilibrium tooth forces. To better characterize the experimental results, the dynamic tooth loads for the experiments of the present work were measured (section 5.2).

The case-carburized gears used for this study have a case and core structure, and therefore the material composition, properties and condition varies with depth from the tooth surface. Both the residual stress profiles and yield strength profiles have been estimated and modeled (sections 5.3 and 5.4).

Most often the pressure distribution of contacting non-conformal surfaces is stated in terms of an idealized Hertzian model where the surfaces are modeled as perfectly smooth, frictionless, and dry. In this work the contact is modeled giving consideration to the lubrication condition. A computer code for the simulation of lubricated line contacts was used in this work. The theoretical basis and numeric procedure for the code that can model either ideally smooth or rough surfaces is described in section 5.5.

The lubricant pressures and tractions acting on the surfaces, as determined by the lubrication analysis, can be used to determine sub-surface contact stresses. A numerical procedure to calculate the stresses induced by the contact loads is presented in section 5.6. In the present work, all stresses are calculated assuming a perfectly elastic response. The stresses induced by load can be superimposed with the residual stress profile to determine the stress tensor at any position within the gear tooth.

To complete the characterization of the experimental conditions of this work, one needs some method for characterizing the intensity of the load as it relates to the state of stress and fatigue damage. The commonly used maximum Hertz pressure provides a simple measure of the load intensity but does not provide for including, directly, the influence of residual stresses nor the influence of roughness. The maximum shear stress and maximum range of shear stresses are two other often used measures of load intensity. More recently, researchers have studied contact problems using ideas and methods from the field of multiaxial fatigue. In the present work, the experimental condition is evaluated in terms of the maximum shear stress, the maximum shear stress range, and a multiaxial fatigue damage index. Procedures to calculate these indices are presented in section 5.7.

Finally, the overall methodology for the evaluation of the experimental condition, as described in detail in the sections to follow, is summarized in section 5.8.

5.2 Measurement of Dynamic Tooth Loads

It is well known that the dynamic gear tooth forces can deviate significantly from the static equilibrium tooth forces. To better characterize the experimental results, the dynamic tooth loads for the experiments of the present work were measured. The technique used was that developed by Rebbechi, Oswald and Townsend (ref. 5.1), and the measurements were done with the assistance of Mr. Oswald. The method of measurement makes use of strategically located strain gages mounted on both sides of adjacent gear teeth (fig. 5.2.1). The strain gages on the front and back sides of the gear teeth exhibit differing responses to normally and tangentially directed forces. A special calibration fixture and procedure was used that permitted the application of only normal forces or combinations of normal and tangential forces at known roll angles of the involute. From the calibration data, one can determine a calibration matrix that relates the measured strains to the applied forces. The instrumented gear was installed into the gear rig used in the present work for fatigue evaluations (fig. 4.2.1), and tooth strains

were measured for operating conditions matching the conditions used for fatigue testing. The calibration matrix was used to calculate the dynamic gear tooth forces from the measured strains.

The dynamic gear tooth force as calculated from the measured strains is provided in figure 5.2.2. The force shown is the force normal to the tooth surface. The solid line depicts the experimental data while the dashed lines are replicates of the data shifted in time by the equivalent of one tooth pitch. The zones of double tooth contact (DTC) and single tooth contact (STC) are illustrated. The mean torque on the gear was calculated from the tooth force data, and the calculated torque value agreed with the known applied torque value to within 1.5 percent. The maximum measured tooth force was 2280 N (513 lb), and the maximum force occurs within the zone of single tooth contact. The nominal contact force for static equilibrium within the single-tooth-contact zone is 1720 N (387 lb). Note that the Hertz stress reported in chapter 4 was based on the nominal contact force for static equilibrium, as has been the custom, but the experimental data here shows that the dynamic factor (about 1.33 times static equilibrium) is significant.

For the purpose of analysis of lubrication and stress, the contact at the low-point of single tooth contact on the driving gear was selected as the critical contact position for analysis. This contact position has been selected by others as the most critical position along the line-of-action (for example Tae, ref. 5.2). Also, from the present investigator's experience this location is the most likely location for pitting to occur for the gear rig used for this work. The measured normal force at this critical contact point was 1960 N (441 lb), and this value was used to complete the analysis presented in chapter 6.

5.3 Modeling of Residual Stress Profiles

Often, the influence of residual stresses that are present in heat-treated gears are not directly considered for the purposes of design or prediction of fatigue life. Instead, the beneficial effects of compressive residual stresses within the case of case-carburized gears are considered by an indirect manner, that is the effect is "built-in" to experimentally determined stress allowable criteria. Such an approach is adequate for sizing gears when using well-established manufacturing methods. However, a more direct treatment is needed to fully understand the influence of surface engineering treatments such as superfinishing, shot-peening, and coatings on gear performance.

It has been established that residual stresses have significant influence on the rolling contact fatigue of bearings. Zaretsky (ref. 5.3) provides a chronological review of relevant research. Recent analytical work by Kotzalas (ref. 5.4) calls for a more complete assessment of residual stress fields for the purpose of bearing fatigue life predictions. The influence of residual stresses on gear fatigue life has also been studied. For example, the beneficial effect of shot-peening on the surface fatigue lives of gears has been explained in terms of the changes in residual stresses induced by the shot-peening operation (refs. 5.5 to 5.7). It should also be considered that the influence of residual stresses on fatigue life may differ for the loads used for accelerated life testing compared to design loads. To best characterize the experimental conditions of the present work, the residual stress profile was modeled and employed for stress analysis.

Residual stress profiles were modeled using data from studies of gear surface fatigue (refs. 5.5 and 5.8). Some of the gears from these studies were made from the same alloy, of the same geometry, and to the same manufacturing specifications as were the gears of the present work. In the previous works, the residual stresses were measured using X-ray diffraction. To obtain residual stress measurements at various depths, material was removed by electropolishing in a sulphuric-phosphoric-chromic acid electrolyte, and the data were corrected for stress relaxation as a result of material removal. The residual stress data for the baseline gears of the previous works were obtained by digitizing the data plotted in the published reports.

The residual stress data of the previous works are collected together and provided in figure 5.3.1. Also depicted by the lines on the plot is a model for the residual stresses that was used for the analysis of chapter 6. The normalized scales provided in figure 5.3.1 are normalized in terms of the maximum Hertzian contact pressure (1.8 GPa) and the Hertzian half-width (241 μm) for the critical contact position as was defined in section 5.2. Note that the residual stress values are significant relative to the stresses induced by load. The residual stress profile model used here is similar to the characteristic residual stress

profile for case-carburized AISI 9310 steel as suggested by Anderson, et al. (ref. 5.9). The directions of the residual stress measurements for the data used here were not reported. However, Batista, et al. (ref. 5.7) in similar work report that the residual stresses in the longitudinal and transverse directions relative to teeth active profiles were quite similar. In the present work the residual stresses depicted in figure 5.3.1 were assumed to apply to both longitudinal and transverse directions.

5.4 Modeling of Yield Strength Profiles

For case-carburized gears as used in this study, the hardness and related mechanical properties vary through the depth of the gear tooth case and core structure. One of the fatigue damage indices selected for the present study (section 5.7) requires an estimate of the yield strength as a function of depth. In this work the yield strength was estimated from measured micro-hardness data (fig. 4.2.3). Since no significant changes were observed to occur due to superfinishing nor due to running of the gears, all of the data shown in figure 4.2.3 were considered as a single dataset for purposes of establishing a characteristic yield strength profile for case-carburized AISI 9310 steel gears. The microhardness data were converted to approximate ultimate tensile strengths using tabulated data for high strength tool steels (ref. 5.10). The uniaxial yield strength was then estimated as 85 percent of the ultimate tensile strength.

The characteristic yield strength profile is depicted by the solid line on figure 5.4.1. The plot also provides the individual point estimates of yield strength calculated from the measured hardness data. The displayed characteristic yield strength profile is defined by the following polynomial,

$$ys = a + b * x + c * x^2 + d * x^3 \quad (5.4.1)$$

where x is the depth below the surface (m) and ys is the yield strength (GPa). The polynomial coefficients used were $a = 1.96$, $b = 533.0$, $c = -1.91 \times 10^6$, and $d = 8.9 \times 10^8$.

5.5 Analysis Method for Determining Contact Pressures of the Lubricated Contact

Often the pressure distribution of contacting non-conformal surfaces is stated in terms of an idealized Hertzian model where the surfaces are modeled as perfectly smooth, dry, and frictionless. However, the experiments described in chapter 4 of the present work, and work by others, have demonstrated that the surface micro-geometries including features on the scale of roughness have a major influence on the fatigue lives of lubricated contacts. Of course, an idealized model using a simplified geometry of smooth surfaces cannot directly account for the influence of surface roughness on fatigue life. Tallian (ref. 5.11) considers that the most far reaching development in the understanding of rolling contact behavior since 1960 has been the recognition of elastohydrodynamic films formed by lubricants in heavily loaded contacts. To best understand the performance of the experimental conditions of chapter 4, the contact conditions were modeled as rough, lubricated surfaces.

The influence of surface roughness on the surface fatigue lives of lubricated contacting surfaces has been established, at least qualitatively, for some time. More than 30 years ago, Smalley, et al., published a state-of-the-art review of the understanding of failure mechanisms of highly loaded rolling and sliding contacts (ref. 5.12). They considered that elastohydrodynamic theory using smooth surface idealizations did provide for reasonable predictions of the steady-state, overall behavior of the contacts. However, they also considered that many important experimentally observed phenomena could be explained only by considering the microtopography of the surfaces and micro-elastohydrodynamics. At that time, micro-elastohydrodynamic theories and models for sliding contacts had only limited development. However, in the three decades that have since passed, micro-elastohydrodynamics theories, methods, and understanding have advanced significantly. In fact, Sayles (ref. 5.13) suggests that in the near future roughness-scale metrology data and contact mechanics will be used not only as research tools but also as production and design tools. The foundations for modern contact mechanics models were comprehensively described by Johnson (ref. 5.14). Later, Liu, Wang, and Lin (ref. 5.15) surveyed the

state-of-the-art for modeling lubricated contacts. Some micro-elastohydrodynamic models have been advanced to include the capability to model time-dependent solutions for real rough surfaces having motion (ref. 5.16). However, there exists a need for further significant research efforts to apply the emerging micro-elastohydrodynamic models to the problems of estimating wear, fatigue damage, and life (ref. 5.17). Considering these facts, the approach adopted in the present work was to model the lubrication condition using a line-contact micro-elastohydrodynamic model having an unchanging real roughness profile. Although such an approach cannot capture certain effects of moving roughness (ref. 5.16) nor certain effects of three-dimensional surface features, it was considered that such detailed treatments of the fluid pressures would not be productive considering the present status of gear fatigue life modeling.

To model the lubrication condition of the gear test conditions used for fatigue testing (chapter 4), the model developed by Cioc, et al. (refs. 5.18 to 5.19) was employed. The model was developed particularly to model the lubrication conditions in helicopter transmissions, and the executable code was available for the present research. The solution for the lubrication condition requires satisfying equations describing mass conservation (Reynolds's equation), energy conservation, geometry within the conjunction including elastic deformations, and the balancing of the external load by the lubricant pressures acting on the surfaces. The Reynolds's equation can be considered as having terms related to the net flows due to pressure gradients (Poiseuille flow), relative surface velocities (Couette flow), and local compressions and expansions (squeeze-film effects). The film thickness is determined by considering the geometries of the undeformed solids, the undeformed surface irregularities, and elastic deformations. In this work, plastic deformations were not considered. This simplification is considered as reasonable since the fatigue lives will be related to the lubrication condition for well run-in surfaces (that is, surfaces with stable geometries). The fluid is modeled as non-Newtonian whereby the energy equation includes a term that is proportional to the product of the shear stress and shear strain rate. The fluid viscosity variation with pressure is modeled with a two-slope exponential model as was proposed by Allen, et al. (ref. 5.20). Fluid density variation with pressure is described by the Dowson-Higginson model (ref. 5.21). Both the fluid viscosity and fluid density depend on the fluid temperature at the inlet, a specified condition by the analyst. In this work isothermal solutions were employed.

The governing equations for the lubrication condition were discretized using 600 equally spaced grid positions over a domain ranging from -1.5 to 1.5 times the Hertzian half-width. The boundary conditions used are a zero gage pressure at the first grid point (before the inlet) and zero gage pressure with zero pressure derivative at the last grid point. The numerical method used is implicit finite difference employing Gauss-Seidel under-relaxation for the Reynolds's and film thickness equations. The model and computer code have been benchmarked against solutions published by other researchers (ref. 5.19).

The lubrication model just described was developed for a cylinder approaching contact with a plane. To apply this model to analyze two contacting spur gears, one can make use of the well known technique to model the contacting teeth using two cylinders, each one having the radius of curvature of the corresponding gear tooth (ref. 5.22). This method to approximate gear contact conditions using two appropriately sized cylinders is used not only for analytical work but also for experimental research (ref. 5.23). For the given position within the mesh cycle, the cylinders and gear teeth have equivalent radii of curvatures, rolling speeds, and sliding speeds at the position of the line contact, figure 5.5.1. The cylinder and plane lubrication model is mathematically equivalent to the two-cylinder model by making use of the appropriate composite radius of curvature.

5.6 Analysis Method for Determining Sub-Surface Stresses

Contact stress analysis can be traced back to the classical work of Hertz (refs. 5.24 and 5.25). Many extensions and applications of the work have been developed. The total solution to the contact stress problem includes both analysis of the deformations and the resulting stresses. In the present work, the solutions for the deformations are included as part of the lubrication (elastohydrodynamic) solution (section 5.5). The lubrication solution provides for surface pressure predictions at discrete grid positions. These pressure predictions can be used to predict the subsurface stress condition at any point within the

solid. Theory of elasticity provides solutions of subsurface stresses for idealized loadings on an elastic half space (ref. 5.14). In the present work, the pressures provided at discrete grid positions (as per the lubrication analysis) were replaced by a triangular shaped pressure distribution extending over two grid-spacings (fig. 5.6.1). One can determine the subsurface stress tensor components due to a triangular shaped pressure distribution using the following set of equations (ref. 5.14),

$$\frac{\sigma_{xx}}{f_0} = \frac{1}{\pi a} \left[(x-a)\theta_1 + (x+a)\theta_2 - 2x\theta + 2z \ln(r_1 r_2 / r^2) \right] - \frac{\mu}{\pi a} \left[(2x) \ln(r_1 r_2 / r^2) + 2a \ln(r_2 / r_1) - 3z(\theta_1 + \theta_2 - 2\theta) \right] \quad (5.6.1)$$

$$\frac{\sigma_{zz}}{f_0} = \frac{1}{\pi a} \left[(x-a)\theta_1 + (x+a)\theta_2 - 2x\theta \right] - \frac{\mu z}{\pi a} \left[(\theta_1 + \theta_2 - 2\theta) \right] \quad (5.6.2)$$

$$\frac{\sigma_{xz}}{f_0} = \frac{-z}{\pi a} \left[\theta_1 + \theta_2 - 2\theta \right] + \frac{\mu}{\pi a} \left[\left[(x-a)\theta_1 + (x+a)\theta_2 - 2x\theta \right] + 2z \ln(r_1 r_2 / r^2) \right] \quad (5.6.3)$$

where r_1 , r_2 , r , θ_1 , θ_2 , and θ are as defined on figure 5.6.1, μ is a proportional friction coefficient that provides for the influence of tangentially directed friction forces, and f_0 is the peak value of the triangular pressure distribution. Equation (5.6.1) here includes a typographical correction to the expression published in the referenced work (change of sign for the second of the square-bracketed terms). The stress state of any position in the solid can be calculated using superposition of stresses due to all pressures at the discrete grid positions. The triangular pressure distribution used here creates an overall pressure distribution that is piece-wise linear.

The method described in the preceding paragraph was implemented as a Fortran subroutine. All floating point calculations were done using double-precision operations. The computer code was verified by reproducing the results of the classical paper by Smith and Liu (ref. 5.26) that describes the influence of tangential and normal loads on contact stresses using Hertzian pressure distribution assumptions. The validated code was applied in the present work (chapter 6) to analyze non-Hertzian pressure distributions.

5.7 Analysis Method for Calculation of Load Intensity and Fatigue Damage Indices

In the previous sections of this chapter, a methodology is presented that permits for the calculation of the state of stress at any desired point in the gear tooth body. To complete the characterization of the experimental conditions of this work, one needs some method for characterizing the intensity of the load as it relates to the state of stress and fatigue damage. The measure of load intensity should be compatible with fatigue mechanism concepts. Ideas concerning the relationship of the lubrication condition to fatigue failure mechanisms were proposed at least 60 years ago (ref. 5.27). Still, contact fatigue failure mechanisms continues to be the subject of experimental research (refs. 5.28 to 5.32). Likewise, research continues on the subject of contact fatigue life models and analysis. Tallian (ref. 5.11) provides a recent, comprehensive review and historical perspective of rolling bearing life models. Kudish and Burris (refs. 5.33 and 5.34) provide another review, and later Kudish proposes a new statistical contact fatigue life model (ref. 5.35). Ai (ref. 5.36) has also proposed a new, novel approach for fatigue life modeling. As is evident by discussions in the literature, the technical community has yet to reach consensus on this topic (refs. 5.11, 5.37, and 5.38). Study of the applications of these recently proposed life models was beyond the scope of the present work. Instead, the experimental conditions of the present work will be expressed in terms of more commonly used stress-based measures of load intensity.

One stress-based measure of load intensity for gear contact fatigue is the reversing shear range on a preselected plane, this plane often referred to as the orthogonal plane. The orthogonal plane for line-

contacts is the X-Z plane of figure 5.6.1. Lundberg and Palmgren, for rolling element bearings (ref. 5.39 and 5.40), made use of the maximum orthogonal reversing shear range as the critical stress causing fatigue. Coy, Townsend, and Zaretsky (refs. 5.41 and 5.42) applied the Lundberg-Palmgren bearing theory to gears. Later, these same authors published experimental work to establish a load-life relationship (ref. 5.43) required for their approach.

A second stress-based index that has been used for roller bearing and gear fatigue is the maximum subsurface shear stress. Coe and Zaretsky made use of maximum subsurface shear to predict the influence of interference fits on roller bearing fatigue life (ref. 5.44). The maximum shear stress has also been used to explain the influence of shot-peening induced residual stresses on gear surface fatigue (refs. 5.5 and 5.6). In general, the maximum subsurface shear stress occurs at a different depth below the surface and on a different plane as compared to the maximum reversing orthogonal shear stress.

Moyar has critiqued the use of single stress type criteria for rolling contact fatigue (ref. 5.45). The previously mentioned maximum shear and maximum reversing orthogonal shear stresses are examples of single stress criteria. He considers that the single stress criteria are inconsistent with established multiaxial fatigue criteria for crack initiation, and he illustrates by example the use of the following multiaxial bulk fatigue initiation criteria,

$$\tau_a \left(1 + \frac{\sigma_{n,MAX}}{\sigma_y} \right) = \tau_e(N_f). \quad (5.7.1)$$

Here, τ_e is considered as an effective shear stress fatigue initiation criteria (or stress based load intensity measure), τ_a is the reversing shear stress on the critical plane, $\sigma_{n,MAX}$ is the maximum value of the normal stress during the stress cycle on the critical plane, and σ_y is the uniaxial yield strength. This criteria has been applied to study rolling contact fatigue of wheel-railway contacts (ref. 5.46). Others have studied the use of various multiaxial fatigue theories for understanding rolling contact fatigue behavior (refs. 5.2, 5.7, 5.32, 5.47, and 5.48).

In the present work, the experimental conditions were evaluated to calculate three stress-based measures of load intensity (that is the maximum shear, maximum reversing shear range, and multiaxial fatigue criteria of equation (5.7.1)). Here, the critical plane was not presumed but instead the stress field was analyzed systematically by recalculating the stress tensor components after successive rotations about three coordinate system axes using angular increments of 3 degrees rotation. All possible permutations of the three successive rotations were inspected covering, in total, 45 degrees rotation about each axis (rotations beyond 45 degrees was not required owing to the symmetry of the stress tensor and transformations). The selection of equation (5.7.1) as the multiaxial fatigue criteria in the present work was made in the spirit of reference 5.45, that is it was selected to study the potential for multiaxial fatigue concepts to improve understanding and predictive capability as applied to rolling contact fatigue. No claim is being made that this multiaxial fatigue criteria is superior to other available multiaxial fatigue criteria.

5.8 Summary of Methodology

In this chapter methodologies for evaluating the experimental condition were developed. To establish the experimental conditions, the dynamic tooth forces were measured. Selected material properties were also measured. These measured data were then used as part of an analytical procedure to evaluate the stresses below the surface of the gear tooth at the selected critical position of the meshing cycle (the low-point of single tooth contact on the driving member). The analytical procedure included a consideration of the rough, lubricated surfaces. The result of the analysis procedure is a calculated stress tensor (included both stresses induced by load and residual stresses of the case-carburized tooth structure). Finally, a procedure is provided to determine critical-plane stress-based measures of load intensity as relates to fatigue damage. The applications of these procedures with results are the subject of chapter 6.

References

- 5.1. Rebbechi, B., Oswald, F., Townsend, D., "Dynamic Measurements of Gear Tooth Friction and Load," *NASA-TM-103281, AVSCOM-TR-90-C-023*, (1990).
- 5.2. Tae, K., "Fatigue of Surface Engineered Steel in Rolling-Sliding Contact," *Ph.D. Thesis, Imperial College of Science, Technology, and Medicine, London, England*, (1999).
- 5.3. Zaretsky, E., "Rolling Bearing and Gear Material," *Tribology for Aerospace Applications—STLE SP-37*, E. Zaretsky, ed., Society of Tribologists and Lubrication Engineers, Park Ridge, IL (1997).
- 5.4. Kotzalas, M., "A Theoretical Study of Residual Stress Effects on Fatigue Life Prediction," *Tribology Transactions*, **44** [4], (2001).
- 5.5. Townsend, D.; Zaretsky, E., "Effect of Shot Peening on Surface Fatigue Life of Carburized and Hardened AISI 9310 Spur Gears," *NASA TP-2047*, (1982).
- 5.6. Townsend, D., "Improvements in Surface Fatigue Life of Hardened Gears by High-Intensity Shot Peening," *NASA TM-105678*, (1992).
- 5.7. Batista, A., et al., "Contact Fatigue of Automotive Gears: Evolution and Effects of Residual Stresses Introduced by Surface Treatments," *Fatigue Fract Engng Mater Struct*, **23**, (2000)
- 5.8. Townsend, D., Turza, A., and Chaplin, M., "The Surface Fatigue Life of Contour Induction Hardened AISI 1552 Gears," *NASA TM-107017, ARL-TR-808*, (1995).
- 5.9. Anderson, W., et al., "Materials and Processing," *STLE Life Factors For Rolling Bearings—STLE SP-34*, E. Zaretsky, ed., Society of Tribologists and Lubrication Engineers, Park Ridge, IL (1992).
- 5.10. "Strength and Hardness in Tool Steels," *TECH-TOPICS Bulletin 112*, Timken Latrobe Steel Company, Latrobe, PA (not dated).
- 5.11. Tallian, T., "The Evolution of Rolling Bearing Life Models," *Society of Tribologists and Lubrication Engineers*, Park Ridge, IL, (2000).
- 5.12. Smalley, A., et al., "Review of Failure Mechanisms in Highly Loaded Rolling and Sliding Contacts," *AD-657337*, Defense Technical Information Center, Ft. Belvoir, VA, (1967).
- 5.13. Sayles, R., "How Two- And Three-Dimensional Surface Metrology Data Are Being Used To Improve The Tribological Performance And Life Of Some Common Machine Elements," *Tribology International* **34** (5), (2001).
- 5.14. Johnson, K., *Contact Mechanics*, Cambridge University Press, (1985).
- 5.15. Liu G., Wang, Q., Lin, C., "A Survey of Current Models for Simulating the Contact Between Rough Surfaces," *Tribology Transactions*, **42** [3], (1999).
- 5.16. Elcoate C., Evans H., Hughes T., and Snidle R., "Thin Film, Time Dependent, Micro-EHL Solutions With Real Surface Roughness" *Proceedings of the Leeds-Lyon Symposium, Lubrication at the Frontier, 1998*, Elsevier, (1999).
- 5.17. Spikes, H., "Tribology Research in the Twenty-First Century," *Tribology International*, **34** (12), (2001).
- 5.18. Cioc, C., Moraru, L., Keith, T. and Kahraman, A., "An Integrated, Physics Based Lubrication Analysis Method for Helicopter Transmission Gear Surfaces with Asperities," *Proceedings of the American Helicopter Society 56th Annual Forum*, American Helicopter Society, (2000).
- 5.19. Cioc, C., "A Deterministic Elastohydrodynamic Lubrication Model of High-Speed Transmission Components," private communications for the Rotorcraft Industry Technology Association, University of Toledo (2002).
- 5.20. Allen, C., Townsend, D., Zaretsky, E., "Elastohydrodynamic Lubrication of a Spinning Ball in a Nonconforming Groove," *ASME Journal of Lubrication Technology*, **92** [1], (1970).
- 5.21. Dowson, D., Higginson, G., *Elasto-Hydrodynamic Lubrication - SI Edition*, Pergamon Press, (1977).
- 5.22. Snidle, R., Evans, H., Alanou, M., "Gears: Elastohydrodynamic Lubrication and Durability," *Proceedings of the Institution of Mechanical Engineers, Part C*, **214**, (2000).

- 5.23]. Phansalkar, S., Kahrahan, A., Anderson, N., "A Comparative Study of EHL Behavior of Smooth and Asperity Contacts of Helical Gears and Rollers," 56th Annual Meeting of the Society of Tribologists and Lubrication Engineers, Orlando, Florida (2001).
- 5.24. Hertz, H., "On the Contact of Elastic Solids (in German)," *J. Reine und Angewandte Mathematik*, **92**, (1882).
- 5.25. Jones, D., Schott, G., Eds., *Miscellaneous Papers of H. Hertz*, Macmillan and Co., London, (1896).
- 5.26. Smith, J., Liu, C., "Stresses Due To Tangential and Normal Loads on an Elastic Solid With Applications to Some Contact Stress Problems," *Trans. ASME, Series E, J. of Applied Mechanics*, **20**, (1953).
- 5.27. Way, S., "How To Reduce Surface Fatigue," *Machine Design*, (1939).
- 5.28. Bingyuan, J., Xitong, Z., Lei, W., "Estimate by Fracture Mechanics Theory for Rolling Contact Fatigue Life and Strength of Case-Hardened Gear Materials With Computer," *Engineering Fracture Mechanics*, **39** [5], (1991).
- 5.29. Webster, M., Norbart, C., "An Experimental Investigation of Micropitting Using a Roller Disk Machine," *Tribology Transactions*, **38** [4], (1995).
- 5.30. Chiu, Y., "The Mechanism of Bearing Surface Fatigue – Experiments and Theories," *Tribology Transactions*, **40** [4], (1997).
- 5.31. Nélías, D., et al., "Role of Inclusions, Surface Roughness and Operating Conditions on Rolling Contact Fatigue," *Trans. Of ASME, Journal of Tribology*, **121**, (1999).
- 5.32. Alfredsson, B., "A Study on Contact Fatigue Mechanisms," *Department of Solid Mechanics, Royal Institute of Technology, Stockholm, Sweden, Doctoral Thesis no. 44*, (2000).
- 5.33. Kudish, I., Burris, K., "Modern State of Experimentation and Modeling in Contact Fatigue Phenomenon: Part I–Contact Fatigue vs. Normal and Tangential Contact and Residual Stresses, Nonmetallic Inclusions and Lubricant Contamination. Crack Initiation and Crack Propagation. Surface and Subsurface Cracks," *Tribology Transactions*, **42** [2], (2000).
- 5.34. Kudish, I., Burris, K., "Modern State of Experimentation and Modeling in Contact Fatigue Phenomenon: Part II–Analysis of the Existing Statistical Mathematical Models of Bearing and Gear Fatigue Life. New Statistical Model of Contact Fatigue," *Tribology Transactions*, **43** [2], (2000).
- 5.35. Kudish, I., "A New Statistical Model of Contact Fatigue," *Tribology Transactions*, **43** [4], (2000).
- 5.36. Ai, X., "Random Surface Roughness on Fatigue Life of a Lubricated Contact," *Trans. Of ASME, Journal of Tribology*, **120**, (1998).
- 5.37. Tallian, T., Hoeprich, M., Kudish, I., "Discussions With Author's Closure – A New Statistical Model of Contact Fatigue," *Tribology Transactions*, **44** [4], (2001).
- 5.38. Tallian, T., "The Use of Microcontact Pressure Statistics in Fatigue Life Modeling of Rolling Contacts," *Tribology Transactions*, **44** [4], (2001).
- 5.39. Lundberg, G., Palmgren, A., "Dynamic Capacity of Rolling Bearings," *Ing. Vetenskaps Akad., Handl.*, **196**, (1947).
- 5.40. Lundberg, G., Palmgren, A., "Dynamic Capacity of Rolling Bearings," *Ing. Vetenskaps Akad., Handl.*, **210**, (1952).
- 5.41. Coy, J., Townsend, D., Zaretsky, E., "Analysis of Dynamic-Capacity of Low-Contact-Ratio Spur Gears Using Lundberg-Palmgren Theory," *NASA TN D-8029* (1975).
- 5.42. Coy, J., Townsend, D., Zaretsky, E., "Dynamic Capacity and Surface Fatigue Life for Spur and Helical Gears," *Journal of Lubrication Technology*, **98** [2], (1976).
- 5.43. Townsend, D., Coy, J., Zaretsky, E., "Experimental and Analytical Load-Life Relation for AISI 9310 Steel Spur Gears," *Trans. Of the ASME, Journal of Mechanical Design*, **100**, (1978).
- 5.44. Coe, H., Zaretsky, E., "Effect Of Interference Fits On Roller Bearing Fatigue Life," *ASLE Transactions*, **30** [2], (1987)
- 5.45. Moyer, G., "A Critique of Rolling Contact Fatigue Criteria for Bearing Rings with Hoop Stresses," *Tribology Transactions*, **38** [2], (1995).

- 5.46. Moyer, G., Stone, D., “An Analysis Of The Thermal Contributions To Railway Wheel Shelling,” *Wear*, **144**, (1991).
- 5.47. Ringsberg, J. et al., “Prediction of Fatigue Crack Initiation For Rolling Contact Fatigue,” *International Journal of Fatigue*, **22**, (2000).
- 5.48. Socie, M., Marquis, G., *Multiaxial Fatigue*, Society of Automotive Engineers, Inc., Warrendale, Pa., (2000).

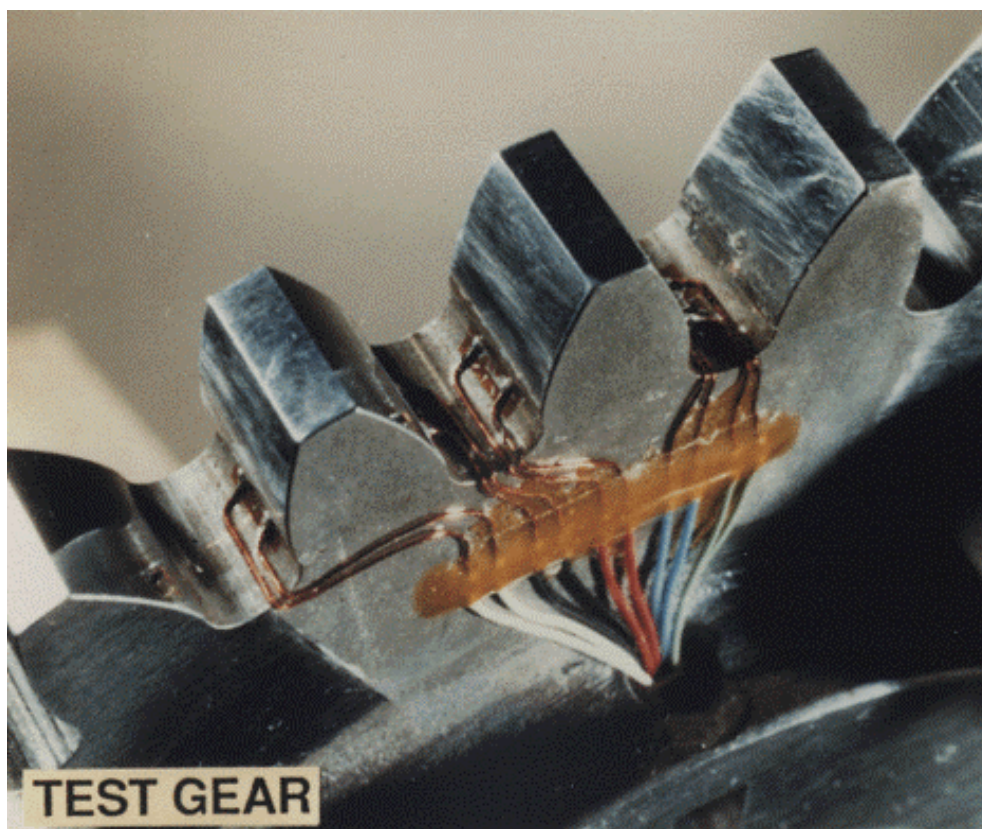


Figure 5.2.1.—Test gear instrumented with strain gages on both sides of successive teeth.

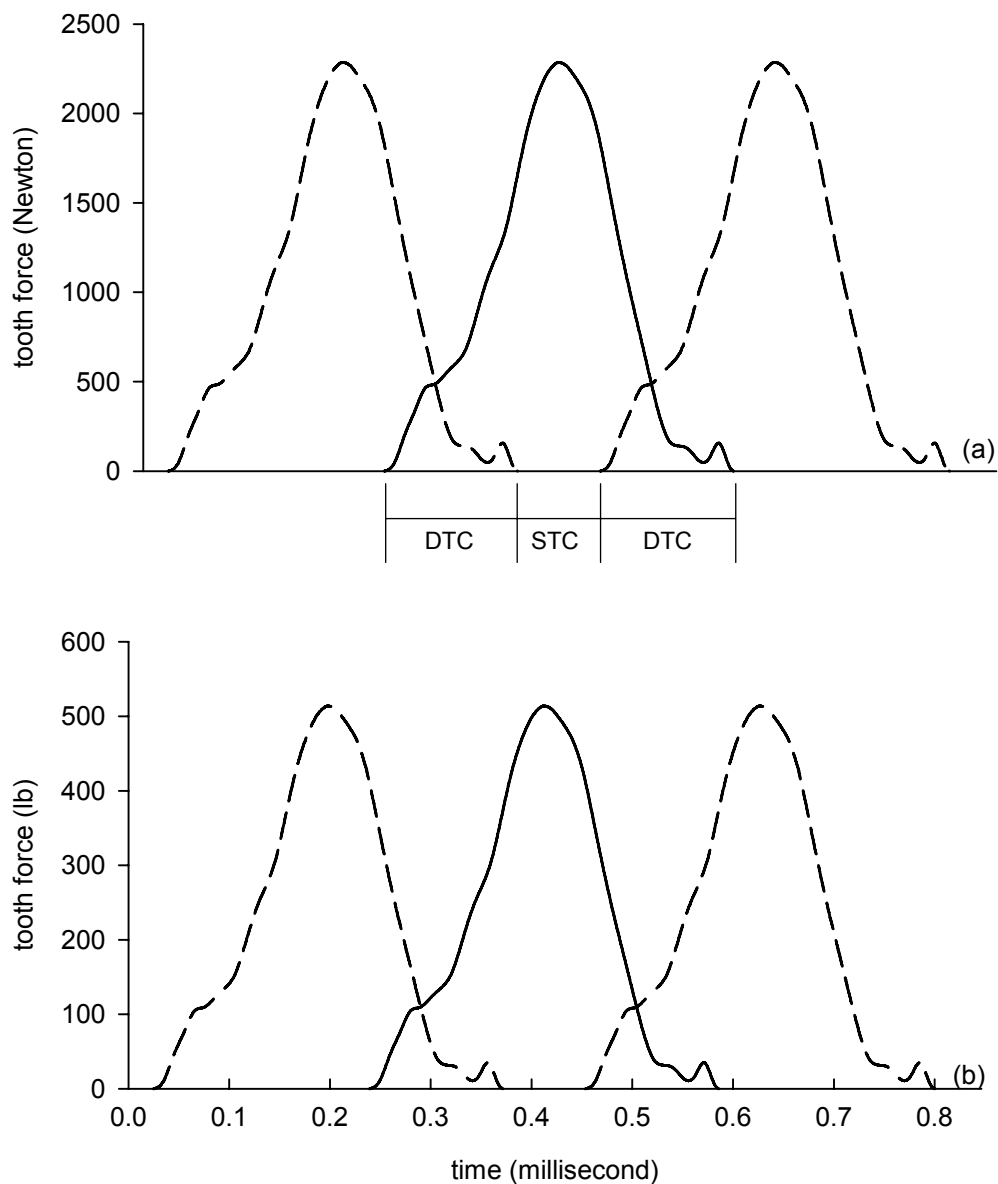


Figure 5.2.2.—Measured dynamic tooth force at the nominal test conditions. The solid line is the measured data, the dashed lines are replicates of the measured data spaced along the ordinate at the equivalent of one tooth pitch. The zones of double tooth contact (DTC) and single tooth contact (STC) are illustrated. (a) Standard International system units. (b) English system units.

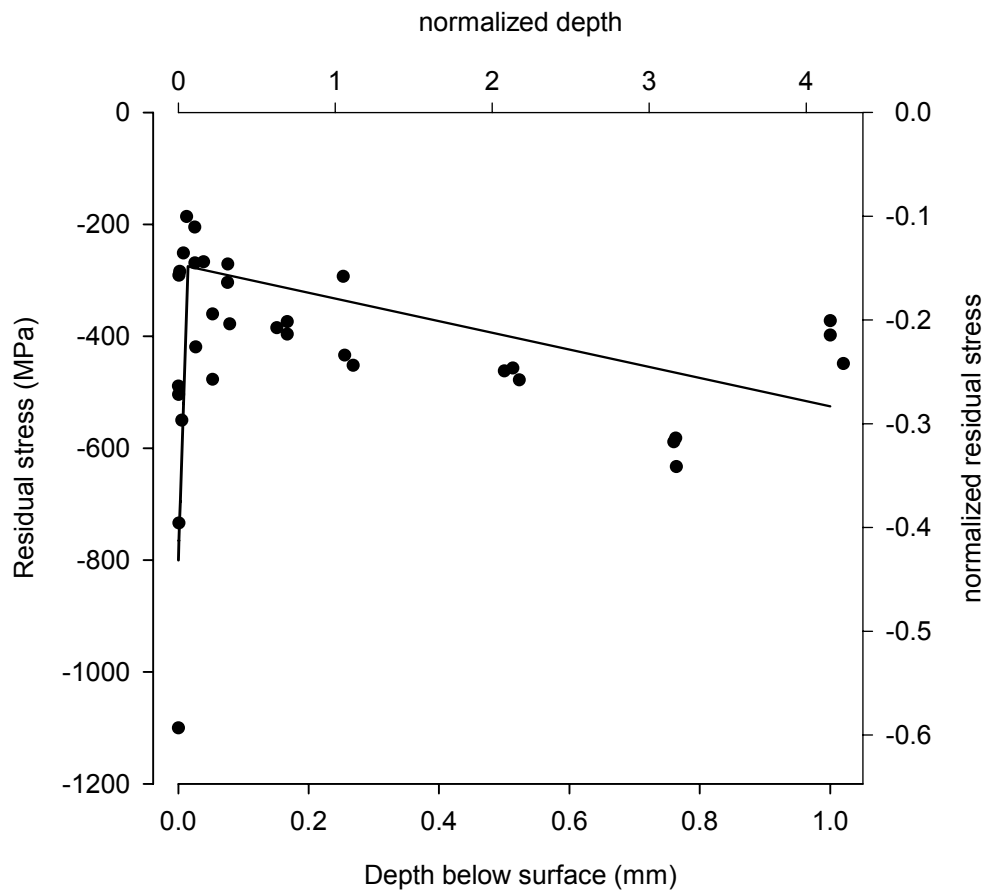


Figure 5.3.1.—Measured residual stress values for several gears and solid line depicting the modeled residual stress profile used for stress analysis (raw data from refs. 5.5 and 5.8).

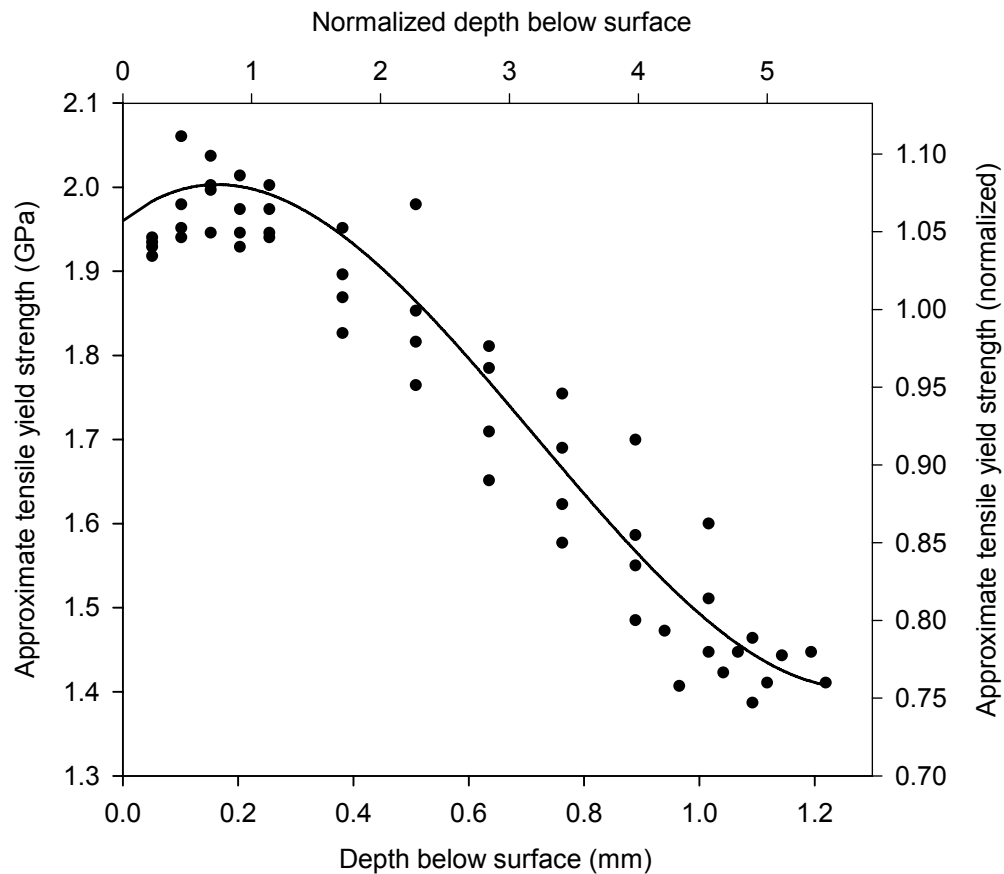


Figure 5.4.1.—Tensile yield strength data points for several gears as based on conversion from measured hardness data (fig. 4.2.3). The solid line depicts the modeled yield strength profile.

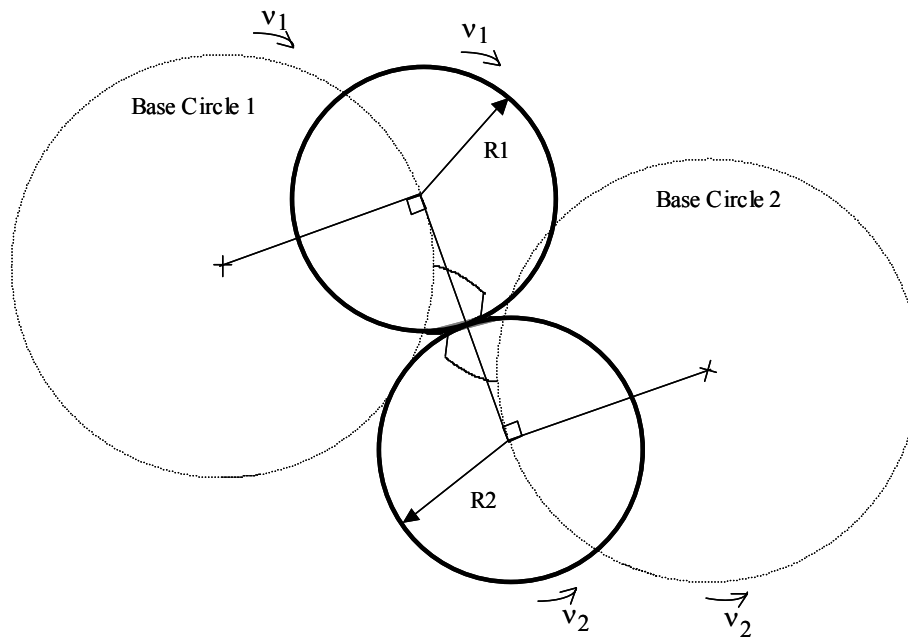


Figure 5.5.1.—Schematic representation of contacting involute spur gear teeth and of a pair of cylinders with equivalent radii of curvatures and equivalent rolling/sliding velocities.

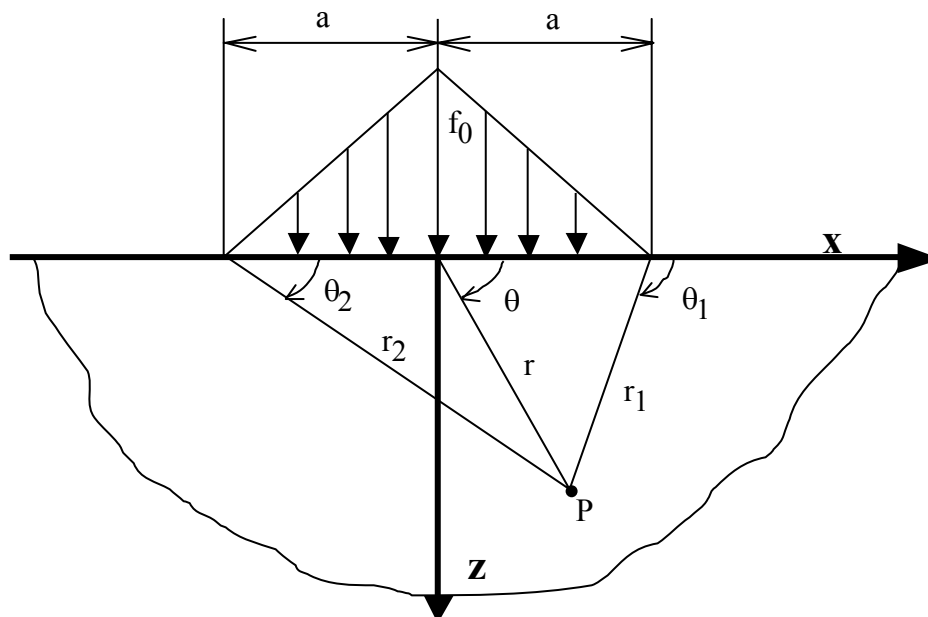


Figure 5.6.1.—Coordinate system and geometry terms used for subsurface stress calculations, in the manner of Johnson, (ref. 5.14).

Chapter 6—Evaluation of the Experimental Conditions (Results)

6.1 Introduction

The experiments documented in chapter 4 showed conclusively that gears with differing as-manufactured surface topographies can have dramatically differing performance characteristics. To apply these laboratory evaluations to products requires engineering understanding, analysis and judgment. In this chapter the methodologies for evaluating the experimental condition (chapter 5) are exercised so as to assist with the application of superfinishing for gears.

The fatigue lives of rolling-and-sliding contacts are often related to surface roughness using the concept of specific film thickness. This concept is an empirical-based function relating the fatigue life to the ratio of the lubricating film thickness to the surface roughness. Section 6.2 provides a further definition of the concept, application of the concept for the present research, and a discussion of its use and limitations.

The experimental conditions of chapter 4 are further characterized using elastohydrodynamics and stress analyses to assess the severity of loading as relates to fatigue. Specifics concerning the application of the methodologies of chapter 5 are provided in sections 6.3 and 6.4. The analysis presented here is, to the author's knowledge, the first evaluation of gear fatigue experiments to include all of the effects of dynamic tooth forces, residual stresses, and yield strength profiles. Analysis are conducted first using a perfectly smooth surface assumption (section 6.5). Such an approach allows for assessing, individually, the influence of friction and residual stress fields. The usual approach of considering only a plane-strain approach was extended in the present research to also include a plane-stress approach.

It has been established by others that superfinished gears operate with reduced tooth friction relative to ground gears. It can be speculated that the reduced friction force is the primary reason for the improved fatigue lives of the superfinished gears. A set of calculations were made to test this speculated relationship. The calculations made use of a smooth-surface model while including the known reduction of friction forces due to superfinishing. Changes in critical-plane stress-based indices were quantified as the ratio of the stress indices for the two cases (a friction coefficient of 0.05 for ground gears and 0.035 for superfinished gears). The ratio of stresses were used to predict a fatigue life ratio. It is shown that such an approach does not offer an adequate explanation for the observed increase in fatigue life.

Since smooth-surface models do not offer an adequate explanation for the improved fatigue lives of superfinished gears, the stress conditions were evaluated using a rough-surface model (section 6.6). Results of the rough-surface modeling provide a qualitative assessment of the stresses as relates to fatigue. The analyses suggest that superfinishing improves fatigue life by influencing the near-surface stress field and, thereby, reducing the rate of the near-surface fatigue process. On the other hand, the stresses at depths corresponding to those of classical sub-surface spalling fatigue analyses are relatively little affected by the roughness features.

6.2 Specific Film Thickness Relation

One approach for applying the results of superfinishing to design practice is to express the surface fatigue lives of gears as a function of the specific film thickness. The specific film thickness is the ratio of the film thickness to the surface roughness. The specific film thickness is sometimes referred to as the "lambda ratio". The film thickness is a calculated central film thickness using a smooth surface model. The roughness is the composite roughness of the two surfaces (square root of the sum-of-the-squares of the roughness root-mean-squared values). Several sets of bearing fatigue life data have been expressed in terms of specific film thickness (refs. 6.1 to 6.5). Townsend and Shimski have expressed gear fatigue data in terms of specific film thickness (ref. 6.6). In their work, the specific film thickness differed from one test group to another owing to differences in oil viscosity and, thereby, differences in the central film thickness.

To be able to compare the data of the present work with the data of the former gearing work (ref. 6.6), the time-to-failure for each test of the former work was determined by digitizing the plotted data. For two of the test groups reported in reference 6.6, testing was done using lubricants that did not include extreme-pressure additives. Fatigue data from testing with these two oils without additives were not included in the present analysis since the present testing did make use of additives. The data were analyzed statistically using the methods of chapter 4 to provide the same mathematical treatment for all data. The measured 10-percent lives along with statistical 90-percent confidence intervals are provided in figure 6.2.1.

Several comments are now provided concerning the specific film thickness to fatigue life relationship for gears as presented in figure 6.2.1. At first glance, it may appear that the data of the present work for superfinished gears are not consistent with that of the former work using ground gears. However, the confidence intervals are large, especially for the larger specific film thickness values. It is certainly possible to draw a curve that is similar in shape to those that have been published for bearings, and the curve would fall within the confidence intervals for all the gear data. Secondly, the gears of the present study were made from a different melt of material than were the gears of the former study. Thirdly, the present author had access to the gears of the former study, and it was noticed that gears from that study, although made from a single melt of steel, were manufactured in two lots. Some of the differences in life among datasets might be attributed to manufacturing lot-to-lot differences. The author declined to draw a proposed curve on figure 6.2.1 since the shape would be speculation. Additional data are required to fully characterize the relationship of specific film thickness to gear surface fatigue life.

The available bearing data that provides the most direct comparison to gear data is that of Skurka (ref. 6.2) for roller bearings. These bearings, like gears, have line contacts (as opposed to elliptical contacts as for ball bearings). Skurka published a life factor curve as a function of specific film thickness. The curve was based on tests of 730 bearings. The published life factor curve demonstrated an increase in bearing fatigue life as specific film thickness was increased from about 1.0 to about 2.5. Little or no increase in fatigue life was observed for increases of specific film thickness beyond the value of about 2.5. Again for Skurka's data, comparing the fatigue lives for specific film thickness of 2.5 relative to a specific film thickness of 1.3, one finds a life ratio of about 3. The gear data provided in figure 6.2.1 are qualitatively consistent with Skurka's results for roller bearings, and the quantitative life ratios for the superfinished relative to ground gears are similar to Skurka's bearing life ratios.

Although the concept of specific film thickness has become a tool for application of elasto-hydrodynamic theory, the concept has some limitations and should be used with care. For example, some have questioned the validity of using un-run values of surface roughness to predict fatigue life since high values of roughness can be reduced by running-in before a small fraction of the lifetime has been expended (ref. 6.4). (chapter 7 provides an assessment of running-in of the gears of the present work.) Chiu (ref. 6.7) points out some limitations of the specific film thickness concept, especially for values less than one as is sometimes realized in practical design. Another discussion of the specific film concept and limitations is provided by Cann, et al. (ref. 6.8). Moyer and Bahney (ref. 6.9) suggest to modify the specific film thickness calculation to provide a consideration of the contact width and the surface wavelengths to be included to quantify roughness. Lubrecht and Venner (ref. 6.10) have proposed a closed-form empirical relation for a modified specific film parameter for rough contacts that is based on the ratio of deformed surface roughness inside the contact divided by the micro-elastohydrodynamic film thickness. In spite of the on-going developments and limitations, the specific film thickness concept can certainly provide guidance for improving the performance of gears and bearings (ref. 6.11).

6.3 Elastohydrodynamic Analysis

To better characterize the experimental conditions of chapter 4, elastohydrodynamic and stress analysis were conducted to assess the severity of loading. Elastohydrodynamic analyses were done to determine the lubricant fluid pressure distributions. The pressure distributions were needed so that surface and subsurface stresses could be evaluated (section 6.4 to follow). The fundamentals of the lubrication

model were described in section 5.5, and the calculations were done using a code that was developed by others (ref. 6.12) and available for the present effort.

The elastohydrodynamic modeling relies on the following assumptions and approximations (see ref. 6.12 for details):

1. continuum mechanics apply, and the bodies and the fluid are homogeneous,
2. the contact is modeled as a line contact (no variations in the direction of the gear tooth face-width), implying no misalignment of the teeth and no crowning of the tooth
3. the bodies deform elastically,
4. the fluid and bodies are all at the same temperature at all locations,
5. the fluid response to shear is non-Newtonian,
6. the fluid viscosity depends on the local pressure as modeled by the modified two-slope Allen approximation (fig. 6.3.1),
7. the fluid density depends on the local pressure as modeled by the Dowson-Higgison approximation (ref. 6.13),
8. the boundary condition for the fluid approaching the conjunction inlet is zero gage pressure at a distance of 1.5 normalized units from the contact center (one normalized distance unit equals the Hertzian contact half-width),
9. the boundary condition for the fluid beyond the outlet is zero gage pressure and zero pressure variation at a distance of 1.5 normalized units from the contact center,
10. the geometry (including roughness) is stationary with respect to load,
11. the involute gear geometry is approximated by two equivalently sized cylinders (fig. 5.5.1),
12. the loads, speeds, and geometries are those for the gears operating at the low point of single-tooth-contact on the driving member.

Numerical inputs for the elastohydrodynamic model are provided in table 6.3.1. The geometry and speeds needed for the lubrication model were determined from the spur gear geometry (table 4.2.2) and the operating speed of 10,000 r.p.m. The equivalent radius and speeds were calculated making use of the equivalent cylinder concept (fig. 5.5.1). The rolling speed is the average of the two tangential speeds, while the slide-to-roll ratio is the ratio of the tangential speed difference divided by the rolling speed. The temperature-dependent lubricant properties were taken to be those for the average of the oil-inlet and oil-outlet temperatures as was observed during the experiments.

6.4 Subsurface Stress Analysis Method and Validation

The pressure distributions as predicted by the elastohydrodynamic analyses were used as input data for stress analysis. The stress analysis method, fully described in chapter 5, is briefly restated here. The elastohydrodynamic analyses provided a set of triangular-shaped, overlapping pressure distributions that, taken together as a set, provided for an overall piecewise-linear pressure distribution. The stress in the body due to any one of the individual triangular-shaped pressure distributions was approximated using the solution for such a distribution on an elastic half-space (section 5.6 and fig. 5.6.1). The stress state was determined by superposition of the stresses due to each individual triangular-shaped distribution plus superposition of the modeled residual stress distribution.

Since the stresses are a result of the pressure distribution predicted by the elastohydrodynamic analysis, all of the assumptions and approximations used (section 6.3) also form part of the set of assumptions for the stress analysis. Additional assumptions needed for the stress analysis are,

1. the contact tractions are a constant coefficient times the normal contact pressures,
2. the material response is either plane stress or plane strain (the assumption for a particular calculation to be made clear in the discussion),
3. since the temperatures of the bodies are assumed constant throughout, thermal stresses are ignored,

4. stresses are due to the contact pressures only, any stresses due to the bending of the gear tooth are ignored.

The assumption stated concerning tractions implies a gross sliding condition. This is appropriate since the selected gear contact position is away from the pitch point, and so the condition of the contact is one of combined rolling and sliding.

The procedure to calculate the state of stress at any particular point was repeated for a rectangular grid of discrete points. Describing the grid in terms of units normalized to the Hertzian half-width, the rectangular grid covered ± 3.0 units along the direction of the tooth involute (x-coordinate) and 2.5 units of depth below the surface (z-coordinate). The grid spacing along the direction of the x-coordinate was 0.2 units outside the contact zone and 0.02 units inside the contact zone. The grid spacing along the z-direction was customized for selected cases of study, but in general the spacing was about 0.02 units from the surface to a depth of 0.1 units, about 0.05 units to a depth of 1.0, and about 0.1 units for further depths. Once the stress field was determined for the entire grid, critical-plane stress-based indices were calculated. The critical-planes were determined using angular increments of 3-degrees for coordinate system rotations. All possible permutations of successive coordinate system rotations were inspected to determine the critical planes.

The stress values and locations are presented using normalized units. Stresses are scaled by the maximum Hertzian contact pressure (1.8 GPa) and linear distances are scaled by the Hertzian half-width (241 μm).

To validate the computerized implementation of the subsurface stress analysis, the present work was compared against previous works. Kannel and Tevaarwerk (ref. 6.14) studied subsurface stresses using numeric solutions, and they provided a table of exact analytic values for a line contact having zero friction and zero residual stresses. Solutions provided by the computer code of the present work compare well with those of this previous work (table 6.4.1) To verify the present computer code for cases of non-zero friction, solutions of the present work were compared to the classical work of Smith and Liu (ref. 6.15). For a friction coefficient of 1/3 and Poisson's ratio of 1/4, both the present code and the previous work predicts the maximum shear stress to occur on the surface, at the location 0.3 normalized units from the contact center, and having the value 0.430 times the maximum pressure of the Hertzian pressure distribution. Since the maximum shear stress is found in a coordinate system orientation other than the one used for initial calculations, this result also validates the computing subroutines that perform tensor transformations and search for maximums to determine planes containing maximum stress-based indices.

6.5 Stress Analysis Results—Smooth Surfaces

Baseline Condition.—To provide a baseline condition for the evaluations of the influences of residual stresses, friction, and roughness, stress calculations were made for an idealized situation representing the experimental operating conditions (chapter 4) while assuming smooth lubricated surfaces, zero friction, zero residual stresses, and a plane strain response. The predicted pressure distribution is provided in figure 6.5.1. The calculated central film thickness is 0.44 μm (17 $\mu\text{in.}$). This is somewhat less than the central film thickness of 0.54 μm (21 $\mu\text{in.}$) as was reported in chapter 4. The minimum film thickness occurs near the exit region of the contact area. The film thickness reported in chapter 4 represents operation at the pitch using the gear tooth load for static equilibrium. The lubricated surface pressure distribution nearly matches the dry Hertzian pressure distribution solution (fig. 6.5.1(b)). From this figure it was anticipated that subsurface stress solutions using this pressure distribution for smooth lubricated surfaces should closely match the solutions for the often-used dry Hertzian pressure distribution solution.

The surface and subsurface stress solutions for the baseline condition are presented as contour plots of selected components of the stress tensor in figures 6.5.2 and 6.5.3. As expected, the stress contours are similar to those for the well-known dry Hertzian solution. The normal stress components (fig. 6.5.2) have distributions that are nearly symmetric relative to the contact centerline. The shear stress contours on two

selected planes (those of maximum shear range and of maximum shear value, fig. 6.5.3) illustrate that the depth of the maximum shear range is shallower than the depth to the maximum shear value.

Another view of the stress state is provided by plotting the stress components for a single depth below the surface while employing stress tensor transformations for coordinate system rotations. Such a view of the stress analysis results are provided in figure 6.5.4 for a subsurface depth of 0.8 normalized units and in figure 6.5.5 for a near surface depth of 0.004 normalized units. The stress components of figure 6.5.4 shows, in the terminology of multiaxial fatigue, out-of-phase loading. The shear stress of figure 6.5.4(d) illustrates the maximum shear range for all depths below the surface and for any possible coordinate system orientation. It is noted that the maximum shear value is not illustrated on figure 6.5.4 since the maximum shear stress is found for a different orientation of the coordinate system. Stress components for the near-surface (depth of 0.004 normalized units) are provided in figure 6.5.5. The stress components in this case are expressed in a coordinate system that illustrates the maximum shear stress value for this particular depth. The stress components of figure 6.5.5 shows, in the terminology of multiaxial fatigue, in-phase loading.

The plots of figures 6.5.4 and 6.5.5 could be thought of as “stress waves” that pass over a point of material as the gears rotate through the meshing cycle, that is the abscissa can be expressed as either a position along the involute for a fixed instant of time or as a time variation for a fixed position along the involute. These “stress waves” can be analyzed to determine stress-based indices for a given depth and coordinate system orientation. As was previously described, in this work the critical planes were not presupposed but instead determined by a search using three-degree increments of rotation for coordinate system orientations. The critical-plane stress-based indices as a function of depth below the surface are provide in figure 6.5.6. It is noted that the individual data points on these plots are not necessarily for the same coordinate system orientation (that is, the critical plane orientation was determined separately for each data point). All three of the indices have similar values at the surface, and the shear range and multiaxial fatigue indices are equivalent for all depths for this baseline condition. The values for the maximum shear value and maximum shear range (and depths of occurrence) match the well-known solutions for a Hertzian pressure distribution.

In this sub-section, baseline stress analyses using simplifying assumptions of zero friction forces and zero residual stresses were presented in the manner of plots of stress contours, plots of “stress waves,” and plots of critical—plane indices. The results match those of the well known Hertzian solution for dry, smooth line contacts. In the sub-sections to follow, the effects of friction and residual stresses are explored.

Influence of Friction Forces.—It has been shown by previous analytical studies that friction forces can significantly alter the subsurface stress field. Some researchers have speculated that the reason why contacts with combined rolling and sliding have significantly smaller fatigue lives as compared to contacts with pure rolling can be explained by the increased friction forces present for sliding conditions (see Chiu, ref. 6.7 for a discussion of this topic). Smith and Liu (ref. 6.15) were the first to conduct a thorough study of the influence of friction forces. Both in that work and the present work, the frictional traction is assumed to be a coefficient times the normal tractions, and Smith and Liu chose a coefficient of 1/3 was included as a limiting case. Although such a large coefficient of friction is not representative for lubricated contacts, such an analysis demonstrates the limiting influence of the frictional forces, and so the friction coefficient value of 1/3 was included as a limiting value for the present work.

The stress contours for the condition matching the baseline case except with the addition of a friction coefficient of 1/3 are provided in figures 6.5.7 and 6.5.8. Concerning the stress contours for the normal stress components (figs. 6.5.2 and 6.5.7), the main effect of friction is to alter the component aligned with the frictional traction. Concerning the stress contours for shear components (figs. 6.5.3 and 6.5.8) the influence is dramatic. The shear stress component XZ for the original coordinate system orientation (fig. 6.5.8(a)) is not the plane containing the maximum shear range, but is provided here to provide a direct comparison to the baseline case (fig. 6.5.3(a)). For the case with friction, the maximum shear range and maximum shear values are now present at the surface rather than subsurface as was the case for zero

friction. The results here match the findings of Smith and Liu (ref. 6.15)], and it can be considered as representative of a dry, unlubricated contact condition.

To demonstrate the influence of friction representative of a lubricated contact, a value of 0.05 for the friction coefficient was selected for study. Calculations were also made for a coefficient value of 0.15 to bridge the value representing the lubricated contact and the dry contact limiting condition (value of 0.33) as was just presented. Results of the calculations are provided as plots of the critical-plane stress-based indices as a function of depth below the surface (fig. 6.5.9). One can see from this plot that although large values of friction have significant influences on the stress-based indices, for a friction coefficient of 0.05 as is representative of a lubricated gear contact (ref. 6.16) the influence is minor. For this studied case using a plain strain response to loads and no residual stresses, the locations of the maximum values of the stress based indices, and their locations, remain essentially unchanged for a friction coefficient range of 0.00 or 0.05. This figure suggests that friction forces do not significantly influence the fatigue lives for well-lubricated contacts. This finding is in agreement with that put forth by Kannel and Tevaarwerk (ref. 6.14). Further discussion of the influence of friction is provided in the latter part of this section.

Influence of Residual Stresses.—The gears of the present study have residual stresses of significant magnitude, largely due to the hardening process. To examine the influence of residual stresses on the stress field and the stress-based indices of fatigue damage, calculations were completed for the case representing the experimental operating conditions (chapter 4) with the assumption of smooth surfaces, a plane-strain response to applied loads, and a friction coefficient of 0.05. Stress contours for two of the normal stress components, both with and without the modeled residual stress field, are provided in figure 6.5.10. It is clearly evident that the residual stress field have a significant influence for the near surface (depths to about 0.05 normalized units). It is also noted that although the patterns of the stress contours at further depths are similar for the cases with and without the residual stresses, the absolute values of the stresses have been affected.

The critical-plane stress-based indices of fatigue damage with residual stresses included are provided in figure 6.5.11 for comparison to figure 6.5.9 (with no residual stresses). As one should expect, the modeled residual stress field that includes non-zero components only for normal stresses does not influence the maximum shear range index (figs. 6.5.9(b) and 6.5.11(b)). The influence of the residual stress field on the multiaxial fatigue parameter (figs. 6.5.9(c) and 6.5.11(c)) is limited to a shallow depth, and for the near surface the added residual stress field produces a somewhat smaller value for the index. Also for the multiaxial fatigue index, the residual stress field has the property of suppressing the influence of the frictional forces on the multiaxial fatigue index.

Concerning the maximum shear values (figs. 6.5.9(a) and 6.5.11(a)), the residual stress field influences this index of fatigue damage for all values of the friction coefficient. Concentrating for a moment on the relationships for depths greater than 0.1 normalized units, the locations and values for the subsurface maximum values has been affected. The residual stress field produces maximum subsurface values at a shallower depth, and the maximum values have been reduced, relative to the zero residual stress solutions. Now giving attention to the very near surface, the added residual stress field acts to increase the maximum shear value at the surface, especially for the larger values of friction coefficient. This observation indicates an increased susceptibility to surface-initiated fatigue owing to increases in compressive residual stresses at the surface, an observation that is at odds with the concept that increased compressive residual stresses will delay or suppress fatigue crack initiation and growth. This apparent contradiction might be due to the simplified treatment of the residual stress field. That is, four of the six independent components of the stress tensor are presumed zero (for lack of better information), but the actual residual stress field likely included non-zero values for some or all of these components.

The influence of residual stresses on the maximum shear stress has been used in previous works to explain the beneficial influences of shot-peening on the surface fatigue lives of gears (refs. 6.17 and 6.18). However, details of the approaches differed for each of the two previous works, and those approaches differs from the one used in the present work. For the older of the two previous works (ref. 6.17), the maximum shear stresses were calculated at a depth of 178 μm (0.007 in.), corresponding to the depth of maximum shear stress for zero residual stresses. For the more recent of the two previous works

(ref. 6.18), the maximum shear stresses were calculated for a depth of 126 μm (0.005 in.). It was stated that the selection of this depth provided a good correlation to the experimentally observed fatigue life ratios using an established stress-life relation. The approaches taken in the present work and the former works differ. In the previous works, a single preselected depth below the surface is used for making comparisons. In the present work, the residual stress field is first superimposed on the load-induced stress field, and only after the superposition is made then the search is undertaken for critical planes and for depths to the maximum values. The new approach shows that the magnitudes and locations (depths) of the critical stress indices are affected by the residual stress fields. Plots such as provided in figure 6.5.11 provide a more complete picture of the influence of residual stress fields as opposed to providing focus to a preselected, single depth below the surface.

As is evident from the preceding discussion, there is yet a fully validated and complete treatment of the influence of compressive residual stresses on the surface fatigue of rolling and sliding contacts. The recent publication of Kotzalas (ref. 6.19) provides still another theoretical treatment employing a discrete volume approach and a Von Mises stress index. He noted the importance of knowing the most accurate values of the residual stress field, and he suggests that the full (6x6) residual stress tensor should be measured and investigated as relates to fatigue life. Although the present author agrees that further investigations of the influence of residual stresses for the purpose of fatigue life prediction is warranted, such investigations were beyond the scope of this project.

Plane Strain vs. Plane Stress Responses.—All calculations and results presented in the preceding sections of this chapter have made use of a plane strain response to the applied loads. This would be considered as a reasonable approximation for evaluating the stress condition at the center of the contact for many practical gearing applications where the tooth face width is much larger than the total tooth height. However, for the gears used in the present work, the tooth width and tooth height are of similar dimensions, and perhaps the usual plane-strain approximation is inadequate. Furthermore, even for the more usual situation of a relatively large face width for the tooth, the stress condition must vary from the plane strain condition near the free edges of the tooth. To further explore the stress condition, the critical-plane stress-based indices of fatigue damage were calculated for a plain stress response to the imposed load. The condition analyzed represents the experimental conditions of chapter 4 with a friction coefficient of 0.05 and a superimposed residual stress field.

Results of the calculations using the plane-stress response to applied load are provided in figure 6.5.12. These results can be directly compared to figure 6.5.11. The relative magnitudes of critical indices at the surface and subsurface differ greatly from the plane-strain to plane-stress responses. The maximum predicted shear stress is greater for the plane-stress response relative to the plane-strain response for all depths. Concerning the shear range and multiaxial parameter indices for friction coefficients less than 0.15, the plane-strain response produces values at the surface that are significantly less than the maximum values that occur in the subsurface. On the other hand, for the plane-stress response the values at the surface and subsurface are of similar magnitude. In general, the results from the plane-stress calculations tend to draw more attention to the stress condition at the surface, while for the plane-strain results one's attention may be drawn to the subsurface as the critical region with the highest indices of fatigue damage. These results, along with the well known deviations of edge stresses for line contacts (ref. 6.20), calls for a full three-dimensional approach as opposed to the commonly used line-contact simplifying assumption as used in the present work. Such a full three-dimensional treatment was beyond the scope of the present project.

Influence of Friction Reduction Due to Superfinishing.—It is well-known that elastohydrodynamic lubricated contacts having combined rolling and sliding motions have shorter fatigue lives relative to contacts having pure rolling motions (refs. 6.7, 6.21 to 6.23). Some authors have speculated that the reduction in fatigue life can be attributed to increases in the friction forces due to the sliding motions (see ref. 6.7 for a discussion). It has also been established by experiments that the superfinishing of gear teeth reduces the gear tooth friction forces by approximately 30 percent (ref. 6.24). This gives rise to the question whether the reduction in friction due to superfinishing can offer an adequate explanation for the observed increase in fatigue lives as was demonstrated by the experiments of chapter 4.

In the preceding paragraphs of this section, critical-plane stress-based indices were calculated for varying friction coefficients using an assumed plane-strain response to load and a simplifying assumption of zero residual stresses. From a qualitative assessment of the results, it was deemed that for friction coefficients representative of lubricated contacts, the influence of friction on the indices of fatigue were minor. In this section, the influence of friction is again assessed in a more thorough manner. The assessment is a quantitative one, and the assessment will include considerations of the residual stress field, plane-strain analyses, and plane-stress analyses. The friction coefficient for the ground gears was taken to be equal to 0.050 while the friction coefficient for the superfinished gears was taken to be equal to 0.035. Surfaces are idealized as smooth surfaces.

A quantitative assessment of the influence of friction was completed by calculating a predicted life improvement ratio. The predicted life ratio is related to the ratio of the stress indices using an inverse power-law relationship and an experimentally determined stress-life exponent of 9. This value for the exponent has been suggested by previous researchers for gear contacts (refs. 6.17 and 6.18). Assessments were made both for the near-surface (normalized depth of 0.004 units) and for the subsurface depths where plots of the indices have a local maximum (figs. 6.5.9, 6.5.11, and 6.5.12).

The critical-plane stress-based indices and resulting life predictions are provided in table 6.5.1. The predicted life ratios range from 1.7 to 1.0 depending on the particulars of the stress index, assumed response to load, and depth location below the surface used for the calculations. The selected multiaxial fatigue index used in this work offered little in the way of additional insight nor improved predictive capability. From the experimental results of chapter 4, the best estimate of the life ratio is a value of about 5.0. Therefore, the improvements in fatigue life due to superfinishing are not adequately explained by analytical models that make use of the simplifying assumption of smooth surfaces. This statement holds true even if the reduction of gear tooth friction due to superfinishing is accounted for in the smooth-surface model.

6.6 Stress Analysis Results—Rough Surfaces

As was demonstrated in the previous section, analytical models that use a simplifying assumption of smooth surfaces fail to fully explain the observed increase in fatigue life due to superfinishing of the gear teeth. To provide a qualitative assessment of the rough-surface stress condition, elastohydrodynamic and stress analyses were completed. The analysis made use of a measured surface profile (fig. 6.6.1). This profile is based on measured roughness profiles of two typical superfinished gear teeth. The measurements were completed after the gears had been subjected to fatigue testing. The teeth selected for inspection did not have any visible signs of fatigue features. Details of the inspection machine and method are found in chapter 7. Raw data from the inspection machine were modified before using the data for rough surface lubrication analyses as follows: (1) a 3-point moving average filter was applied to remove spurious data owing to limitations of the measurement method, (2) long wavelength features, greater than about one-half the Hertzian half-width, were removed, and (3) two modified profiles were superimposed to provide a composite profile of two surfaces. The resulting roughness profile was scaled to provide three profiles with varying degrees of roughness. The profile depicted in figure 6.6.1 represents the largest magnitude of roughness analyzed. Note that the scales used to depict the profiles (vertical scale approximately 1000X of the horizontal scale) offer a distorted view of the roughness features. The true slopes of the roughness features are shallow. Figure 6.6.2 offers a view of a small portion of the total profile showing the roughness with a true aspect ratio.

Film thickness prediction from the rough-surface elastohydrodynamic analyses for three scaled-versions of the roughness profile are depicted in figure 6.6.3. For the case of roughness with the largest amplitude (fig. 6.6.3(c)), the undeformed peak-to-valley magnitude equaled about 85 percent of the minimum film thickness for smooth surfaces. One can see that for these operating conditions, the surface roughness features have been significantly flattened inside the contact, and the film thickness does not deviate significantly from a smooth-surface condition. The significant separation of surfaces (relative to

the deformed roughness) would suggest nominally zero wear, and this suggestion is consistent with the negligible change in appearance that occurred when testing the superfinished gears.

The significant deformation of the roughness features suggest that the pressure distribution deviates significantly from the smooth surface pressure profile (else the predicted film thickness profile would more resemble the undeformed roughness). The pressure distribution predictions are provided in figure 6.6.4. These data show that even for near-mirror quality superfinished surfaces, the peak fluid pressures significantly exceed those for the ideally-smooth analysis. For this particular roughness profile and operating condition, the peak pressure exceeds the smooth-surface maximum pressure by about 50 percent.

The pressure distributions were used to analyze the subsurface stress condition, and the results will now be presented for the two scaled profiles of larger magnitudes along with the solution for perfectly smooth surfaces. The stress calculations were made assuming a coefficient of friction of 0.05, a plane strain response to the applied loads, and a superimposed residual stress field per figure 5.3.1.

Stress contour plots for the two of the normal stress components (the 'XX' component aligned with the tooth involute and the 'ZZ' component normal to the tooth surface) are provided in figures 6.6.5 and 6.6.6. These plots show that the influence of the roughness is rather shallow with the most significant roughness features occurring for depths less than 0.1 of the Hertzian half-width. As the roughness increases, the maximum stress values in the very near surface increases, and the influence of roughness extends to increasing depths.

Stress contour plots for the shear stress components 'XZ' are provided in figure 6.6.7. This plane orientation contains the maximum shear-range for the smooth-surface analysis. This plot, like those for the normal stresses, shows that the influence of the roughness is rather shallow. Several significant shear stress reversals occur in the very near surface rather than the single stress reversal as occurs for the smooth surface analysis. Tae (ref. 6.21) made a similar observation and put forth a proposal to consider such fluctuations as multiple stress cycles per passing of an area through a contact. Tae applied a rain flow counting method to analyze such a complicated stress fluctuation.

Another view of the shear stress condition is provided in the plot of figure 6.6.8. The stress tensor data was transformed by a 45 degree rotation about the 'X' axis to provide this view of the stress condition. For the smooth surface case, this plane orientation contains the maximum absolute value of shear stress. A very shallow depth is provided on the ordinate to show the surface details. This view of the data suggests only a very shallow influence of the roughness. The stress condition for rough surfaces is a complicated one, and differing single views of the data may tend to lead one to differing conclusions concerning the relative effects of the roughness.

Attempts were made to analyze the elastohydrodynamic condition for a roughness profile corresponding to a composite profile for run-in, ground gear teeth. Solutions could not be obtained due to limitations of the computer code available for the present research. Modification of the computer code to robustly handle such roughness profiles was beyond the scope of this project.

Discussion.—The elastohydrodynamic and stress analyses presented here offers significant insights concerning the improved fatigue lives for superfinished gears relative to ground gears. The results indicate that even if the undeformed roughness is less than the film thickness, still the roughness deforms significantly as the roughness passes into the contact. The interacting asperities act as locally converging regions and thereby produce pressure ripples. The pressure ripples produce corresponding variations of the stress field in the near surface. While prevailing gear surface fatigue life models are developed on the notion of initiation of spalls at relatively deep locations (ref. 6.25), the analyses conducted here suggest that the most significant influence of the roughness is for a relatively shallow layer of material. Therefore, current gear surface fatigue life models cannot account for the influence of roughness except by the use of empirical-based life adjustment factors.

The present research suggests that the superfinished gears had longer fatigue lives relative to ground gears largely due to a reduced rate of micropitting fatigue processes (surface or near-surface initiated fatigue) rather than due to any reduced rate of spalling (sub-surface) fatigue processes. This finding is in agreement with many analytical studies in the literature for bearings. However, a widely accepted

quantitative fatigue life model for gears that can capture roughness effects directly as part of the stress and fatigue calculations is not available. Several differing approaches that might be adopted have been put forth in the literature for rolling and sliding contacts. References 6.26 and 6.27 provide significant discussions and guidance for ongoing research. Recently published results (ref. 6.28) suggest that a life model for gears should include consideration of the combined bending and contact stresses along with inelastic (shakedown) phenomena. Bending stresses and inelastic phenomena were not included in the present research. Thermal stresses, also neglected in the current effort, might also play a role in the fatigue of gears, and some tools for handling such thermal stresses are being published (refs. 6.29 and 6.30).

New and successful gear surface fatigue life models to guide the application of emerging surface engineering techniques such as superfinishing, coatings, duplex heat treating, and peening operations will likely need to incorporate new findings from several fields of research including metrology, surface statistics, contact mechanics, elastohydrodynamics, tribology, and multiaxial fatigue. A new and widely accepted fatigue-life model for gears will require significant efforts to produce experimental data for validation. Tallian (ref. 6.26) has noted the lack of sufficient experimental data for bearings for validating life models. Especially lacking is data published openly and with sufficient detail of documentation. Such data for gears is even more scarce, and so further experimental investigations of gear fatigue life is recommended to help guide and validate future gear life modeling efforts.

6.7 Conclusions and Recommendations

In this chapter, the experimental conditions used for testing in chapter 4 were evaluated using several approaches. First, the present research on the relation of surface roughness to fatigue life was expressed in terms of specific film thickness. Data for gears are compared to that for bearings, and comments are provided concerning the limitations and use of the specific film concept. Secondly, a series of analysis were made using the simplifying assumption of smooth surfaces. It is demonstrated that such models do not adequately explain the experimentally observed improvements in fatigue life due to superfinishing of gears. Lastly, the stress condition is analyzed using a rough surface model. These rough-surface analyses provide qualitative explanations for the improved fatigue lives. The following specific results were obtained.

1. Expressing available gear fatigue data in terms of specific film thickness, it is possible to draw a curve that is similar in shape to those that have been published for bearings, and the curve would fall within the confidence intervals for all the gear data.
2. Expressing available gear fatigue data in terms of specific film thickness, the quantitative life ratios for the superfinished and ground gears are similar to openly published data for roller bearing life ratios.
3. The present research shows that the magnitudes and locations (depths) of the critical stress indices are affected by residual stress fields.
4. Calculations of the stress condition using a smooth surface model and both plane-strain and plane-stress responses to the applied loads showed differing characteristics. In general, the results from the plane-stress calculations tend to draw more attention to the condition at the surface, while for the plane-strain results one's attention may be drawn to the subsurface as the critical region with the highest indices of fatigue damage.
5. The selected multiaxial fatigue index used in this work to characterize smooth-surface stress analyses offered little in the way of additional insight nor improved predictive capability. This suggests that the changes in friction forces provided by superfinishing does not significantly alter the fatigue phenomena.
6. Qualitative trends from smooth-surface stress analyses suggest that friction forces do not significantly influence fatigue life. This conclusion was reached employing proportional friction coefficients having values representative of well-lubricated contacts.

7. Quantitative predictions of increased fatigue life due to superfinishing are not adequately explained by analytical models that make use of smooth surface models even if the reduction of gear tooth friction due to superfinishing is accounted for in the calculations.
8. Elastohydrodynamic analyses using rough surfaces predict that even for near-mirror quality superfinished surfaces, the peak fluid pressures significantly exceed those for the ideally-smooth analysis. Peak pressure exceeded the smooth-surface maximum pressures by nearly 50 percent for the conditions analyzed herein.
9. Even if the undeformed roughness is less than the film thickness, still the roughness deforms significantly as the roughness passes into the contact.
10. The present research suggests that the superfinished gears had longer fatigue lives relative to ground gears largely due to a reduced rate of micropitting fatigue processes (surface or near-surface initiated fatigue) rather than due to any reduced rates of spalling (sub-surface) fatigue processes.

The following recommendations are made concerning applications of the present work and concerning extensions of this research.

1. The specific film thickness concept can provide guidance for improving the performance of gears. The concept is a useful tool for applying, to practical designs, superfinishing to improve gear fatigue life as was observed in the laboratory evaluations conducted in the present research.
2. Further investigations of the influence of residual stresses for the purpose of fatigue life prediction is warranted.
3. Future modelling efforts should consider a full three-dimensional approach for the stress condition as opposed to the commonly used line-contact simplifying assumption as was used in the present work.
4. New and successful gear surface fatigue life models will likely need to incorporate new findings from several fields of research including metrology, surface statistics, contact mechanics, elastohydrodynamics, tribology, and multiaxial fatigue.
5. Further experimental investigations of gear fatigue life is recommended to help guide and validate future gear life modeling efforts.

Future analytical investigations should give consideration to recently published results concerning the importance of bending stresses, inelastic phenomena, and thermal stresses on gear fatigue performance.

References

- 6.1. Tallian, T., "On Competing Failure Modes in Rolling Contact," *ASLE Transactions*, **10** [4], (1967).
- 6.2. Skurka, J., "Elastohydrodynamic Lubrication of Roller Bearings," *Trans. Of the ASME, Journal of Lubrication Technology*, (1970).
- 6.3. Anderson, W., "Elastohydrodynamic Lubrication Theory as a Design Parameter for Rolling Element Bearings," *ASME Paper 70-DE-19, ASME Design Engineering Conference & Show, Chicago, IL*, ASME, New York, (1970).
- 6.4. Liu, J, Tallian, T., McCool, J., "Dependence of Bearing Fatigue Life on Film Thickness to Surface Roughness Ratio," *ASLE Transactions*, **18** [2], (1975).
- 6.5. *Life Adjustment Factors for Ball and Roller Bearings, An Engineering Design Guide*, ASME, New York, (1971).
- 6.6. Townsend, D.; Shimski, J., "Evaluation of the EHL Film Thickness and Extreme Pressure Additives on Gear Surface Fatigue Life," *NASA TM-106663*, (1994).
- 6.7. Chiu, Y., "The Mechanism of Bearing Surface Fatigue – Experiments and Theories," *Tribology Transactions*, **40** [4], (1997).

- 6.8. Cann P., Ioannides E., Jacobson B., Lubrecht A., "The Lambda Ratio - A Critical Reexamination," *Wear*, **175** [1-2], (1994)
- 6.9. Moyer, C., Bahney, L., "Modifying the Lambda Ratio to Functional Line Contacts," *Tribology Transactions*, **33** [4], (1990).
- 6.10. Lubrecht, A., Venner, C., "Elastohydrodynamic Lubrication of Rough Surfaces," *Proc. Instn. Mech. Engrs., Part J*, **213**, (1999).
- 6.11. Moyer, C., "Using the Modified Lambda Ratio to Advance Bearing and Gear Performance," *SAE Technical Paper Series 901625, presented at the International Off-Highway & Powerplant Congress and Exposition*, SAE International, Warrendale, PA (1990).
- 6.12. Cioc, C., Moraru, L., Keith, T. and Kahraman, A., "An Integrated, Physics Based Lubrication Analysis Method for Helicopter Transmission Gear Surfaces with Asperities," *Proceedings of the American Helicopter Society 56th Annual Forum*, American Helicopter Society, (2000).
- 6.13. Dowson, D., Higginson, G., *Elasto-Hydrodynamic Lubrication - SI Edition*, Pergamon Press, (1977).
- 6.14. Kannel, J., Tevaarwerk, J., "Subsurface Stress Evaluations Under Rolling/Sliding Contacts," *Trans. Of the ASME, J. of Tribology*, **106**, (1984).
- 6.15. Smith, J., Liu, C., "Stresses Due To Tangential and Normal Loads on an Elastic Solid With Applications to Some Contact Stress Problems," *Trans. ASME, Series E, J. of Applied Mechanics*, **20**, (1953).
- 6.16. Rebbechi, B., Oswald, F., Townsend, D., "Dynamic Measurements of Gear Tooth Friction and Load," NASA-TM-103281, AVSCOM-TR-90-C-023, (1990).
- 6.17. Townsend, D.; Zaretsky, E., "Effect of Shot Peening on Surface Fatigue Life of Carburized and Hardened AISI 9310 Spur Gears," NASA TP-2047, (1982).
- 6.18. Townsend, D., "Improvements in Surface Fatigue Life of Hardened Gears by High-Intensity Shot Peening," NASA TM-105678, (1992).
- 6.19. Kotzalas, M., "A Theoretical Study of Residual Stress Effects on Fatigue Life Prediction," *Tribology Transactions*, **44** [4], (2001).
- 6.20. Kannel, J., Hartnett, M. "Theoretical And Experimental Evaluation Of Edge Stresses Under Severe Edge Loads," *ASLE Transactions*, **26** [1], (1983).
- 6.21. Tae, K., "Fatigue of Surface Engineered Steel in Rolling-Sliding Contact," *Ph.D. Thesis, Imperial College of Science, Technology, and Medicine, London, England*, (1999).
- 6.22. Graham, R., Olver, A., Macpherson, P., "An Investigation into the Mechanisms of Pitting in High-Hardness Carburized Steels," *ASME Paper 80-C2-DET*, ASME, New York, (1980).
- 6.23. Webster, M., Nobart, C., "An Experimentel Investigation of Micropitting Using a Roller Disc Machine," *Tribology Transactions*, **38** [4], (1995).
- 6.24. Snidle, R.; Evans, H.; Alanou, M., "The Effect of Superfinishing on Gear Tooth Profile," *AD-A327916*, Defense Technical Information Center, Ft. Belvoir, VA, (1997). Also available from NASA Scientific and Technical Information Center, <http://www.sti.nasa.gov/> accessed October 3, 2000.
- 6.25. Coy, J., Townsend, D., Zaretsky, E., "Analysis of Dynamic-Capacity of Low-Contact-Ratio Spur Gears Using Lundberg-Palmgren Theory," *NASA TN D-8029*, (1975).
- 6.26. Tallian, T., "The Evolution of Rolling Bearing Life Models," *Society of Tribologists and Lubrication Engineers*, Park Ridge, IL, (2000).
- 6.27. Tallian, T., "The Use of Microcontact Pressure Statistics in Fatigue Life Modeling of Rolling Contacts," *Tribology Transactions*, **44** [4], (2001).
- 6.28. Kapoor, A., Morales-Espejel, G., Olver, A., "A Shakedown Analysis of Simple Spur Gears," *Tribology Transactions*, **45** [1], (2002).
- 6.29. Handschuh, R., "Thermal Behavior of Spiral Bevel Gears," *Ph.D. Thesis, Case Western Reserve University, Cleveland, Ohio*, (1993).
- 6.30. Liu, G., Wang, Q., "Thermoelastic Asperity Contacts, Frictional Shear, and Parameter Correlations," *Trans. Of the ASME, Journal of Tribology*, **122**, (2000).

TABLE 6.3.1.—LUBRICATION MODEL PARAMETERS TO SIMULATE THE
EXPERIMENTAL OPERATING CONDITIONS AT THE LOW-POINT OF
SINGLE-TOOTH-CONTACT ON THE DRIVING GEAR.

<u>Geometry and speeds</u>	
Rolling speed	15.92 (m/s)
Slide/roll ratio	0.223
Equivalent radius	0.0075 m
<u>Fluid data</u>	
Viscosity	0.01703 Pa-s
Density	866 kg/m ³
Viscosity variation with pressure	Allen model (fig. 6.3.1)
Load	701500 N/m

TABLE 6.4.1.—COMPARISON OF RESULTS FOR THE NUMERIC PROCEDURE AND FOR EXACT
ANALYTICAL SOLUTION USING A HERTZIAN PRESSURE DISTRIBUTION AND ZERO FRICTION,
FOR A DEPTH OF ONE-HALF THE HERTZIAN HALF-WIDTH (EXACT RESULTS PER REF. 6.14)

Location	Sxx / Pmax		Syy / Pmax		Szz / Pmax		Sxz / Pmax	
x / b	numeric	exact	numeric	exact	numeric	exact	numeric	exact
-0.8	-0.299	-0.299	-0.227	-0.227	-0.498	-0.500	+0.247	+0.247
-0.5	-0.313	-0.313	-0.302	-0.301	-0.745	-0.745	+0.176	+0.176
-0.3	-0.330	-0.329	-0.334	-0.334	-0.842	-0.842	+0.107	+0.107
0.0	-0.342	-0.342	-0.352	-0.352	-0.894	-0.894	0.000	0.000
0.3	-0.329	-0.329	-0.334	-0.334	-0.842	-0.842	-0.107	-0.107
0.5	-0.313	-0.313	-0.302	-0.301	-0.745	-0.745	-0.176	-0.176
0.8	-0.299	-0.299	-0.227	-0.227	-0.498	-0.500	-0.247	-0.247

TABLE 6.5.1.—PREDICTED FATIGUE LIFE RATIOS DUE TO REDUCTIONS OF FRICTION OWING TO SUPERFINISHING (AS DETERMINED USING CRITICAL-PLANE STRESS-BASED PARAMETERS FOR SMOOTH SURFACES MODELING.)

PARAMETERS FOR SMOOTH SURFACES MODELING				
Assumed response to load	Depth below surface ¹	Friction coefficient	Stress index value ²	Predicted life ratio
Maximum shear stress index ³				
Plane strain	0.004	0.035 (superfinish)	0.2623	1.7
Plane strain	0.004	0.050 (ground)	0.2781	
Plane stress	0.004	0.035 (superfinish)	0.4954	1.1
Plane stress	0.004	0.050 (ground)	0.4986	
Plane strain	0.450	0.035 (superfinish)	0.2631	1.2
Plane strain	0.450	0.050 (ground)	0.2678	
Shear range index				
Plane strain	0.004	0.035 (superfinish)	0.2397	1.4
Plane strain	0.004	0.050 (ground)	0.2496	
Plane stress	0.004	0.035 (superfinish)	0.4995	1.2
Plane stress	0.004	0.050 (ground)	0.5217	
Plane strain	0.500	0.035 (superfinish)	0.4954	1.0
Plane strain	0.500	0.050 (ground)	0.4954	
Plane stress	0.500	0.035 (superfinish)	0.4947	1.0
Plane stress	0.500	0.050 (ground)	0.4947	
Multiaxial fatigue index				
Plane strain	0.004	0.035 (superfinish)	0.1926	1.3
Plane strain	0.004	0.050 (ground)	0.1958	
Plane stress	0.004	0.035 (superfinish)	0.4130	1.0
Plane stress	0.004	0.050 (ground)	0.4130	
Plane strain	0.500	0.035 (superfinish)	0.4947	1.0
Plane strain	0.500	0.050 (ground)	0.4947	
Plane stress	0.500	0.035 (superfinish)	0.4947	1.0
Plane stress	0.500	0.050 (ground)	0.4947	

Notes: (1) Depth below surface is normalized to Hertzian half-width of 241 μm .
(2) Stresses are normalized to maximum Hertzian pressure of 1.8 GPa.
(3) Maximum shear stress index does not have a local maximum in the subsurface for the plane-stress response.

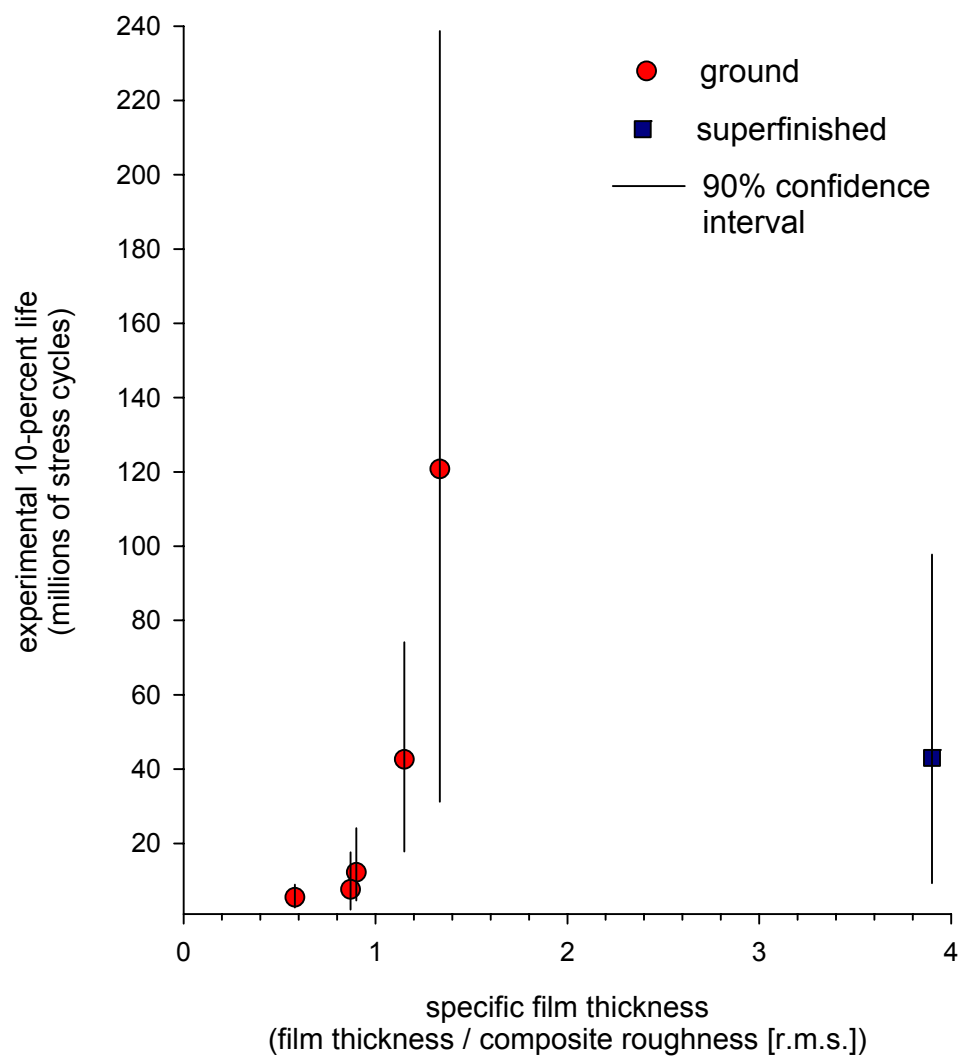


Figure 6.2.1.—Experimental 10-percent lives as a function of specific film thickness with 90-percent pointwise confidence bands. Data for the superfinished gears are from the present study. Data for the ground gears taken from ref. 6.6.

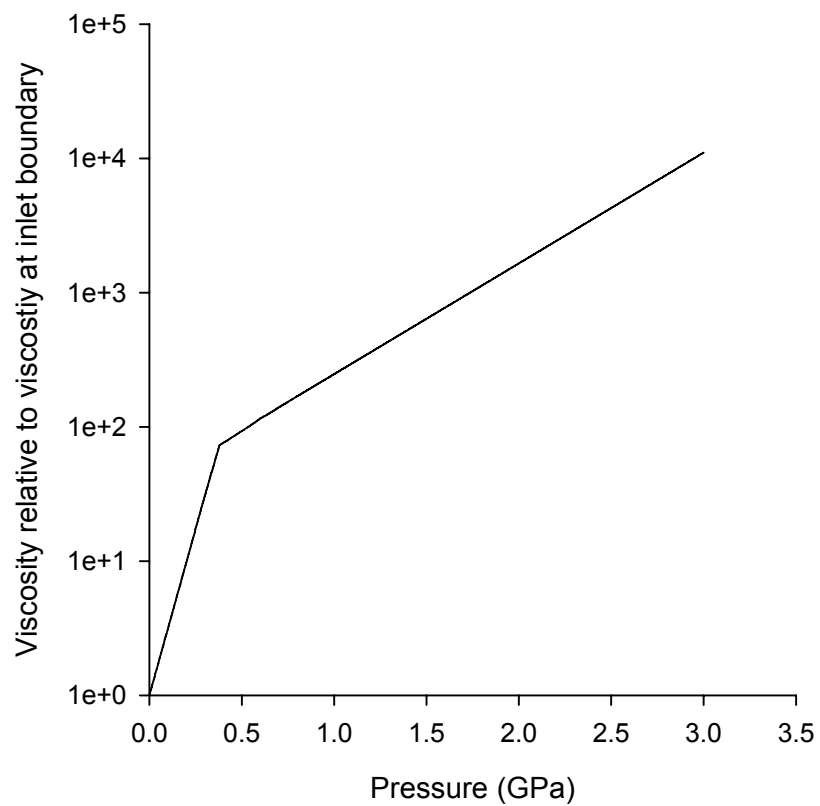


Figure 6.3.1.—Two-slope model used for pressure-viscosity relation. The computerized implementation includes a modification to this two-slope model, to provide for a continuous derivative, using a third-degree polynomial (ref. 6.12).

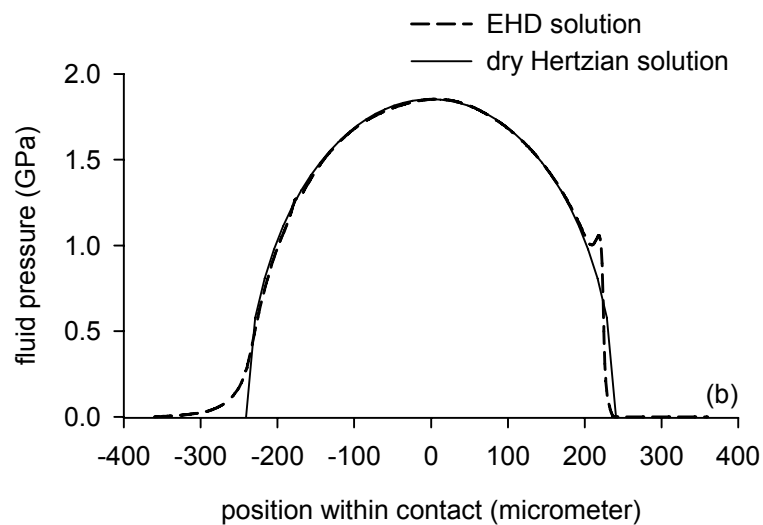
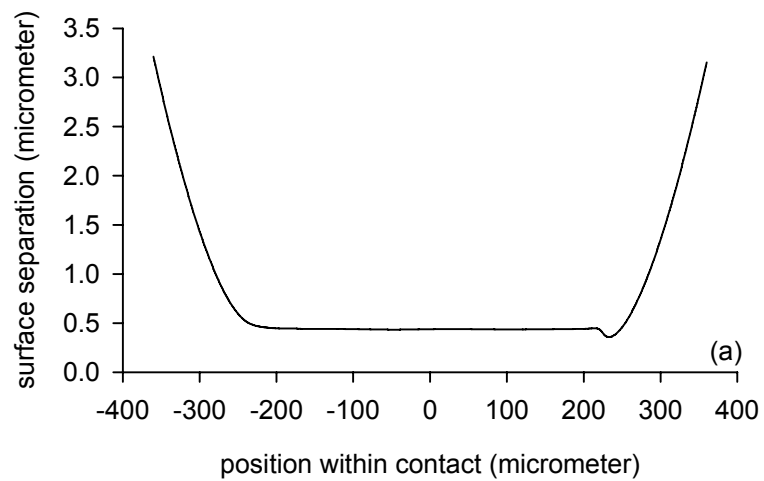


Figure 6.5.1.—Predicted lubrication condition using smooth surface elastohydrodynamic (EHD) model. (a) Surface separation. (b) Fluid pressures with comparison to the dry Hertzian solution for the line-contact.

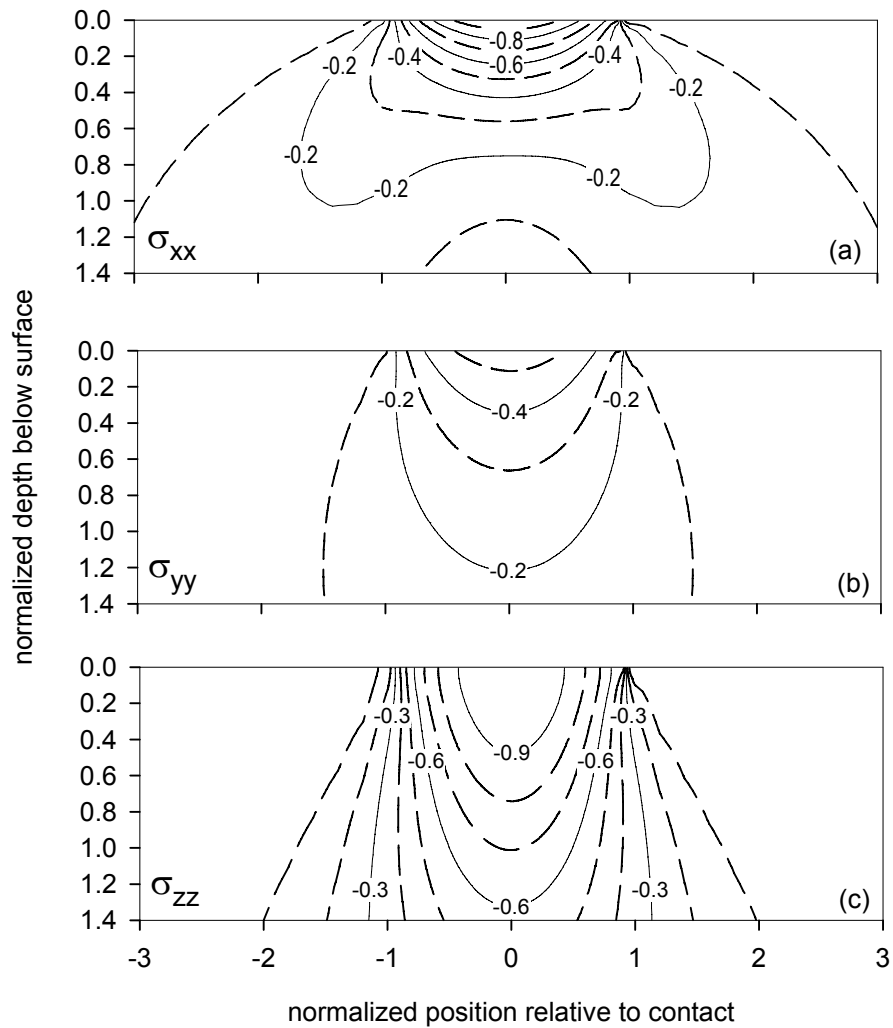


Figure 6.5.2.—Contour plots of normal stress components for the baseline case of smooth lubricated surfaces, zero friction forces, and zero residual stresses. (a) Component in the direction of the gear tooth involute. (b) Component in the direction of the gear tooth face width. (c) Component in the direction of depth below the surface.

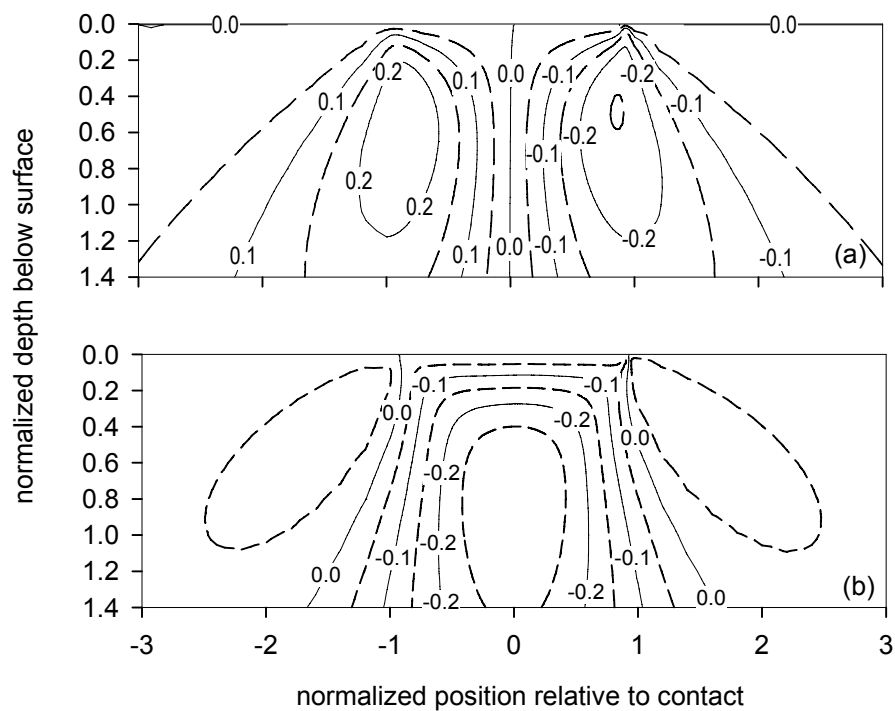


Figure 6.5.3.—Contour plots of shear stress components for the baseline case of smooth lubricated surfaces, zero friction forces, and zero residual stresses. (a) Shear component on the plane of maximum reversing shear. (b) Shear component on the plane of maximum shear value.

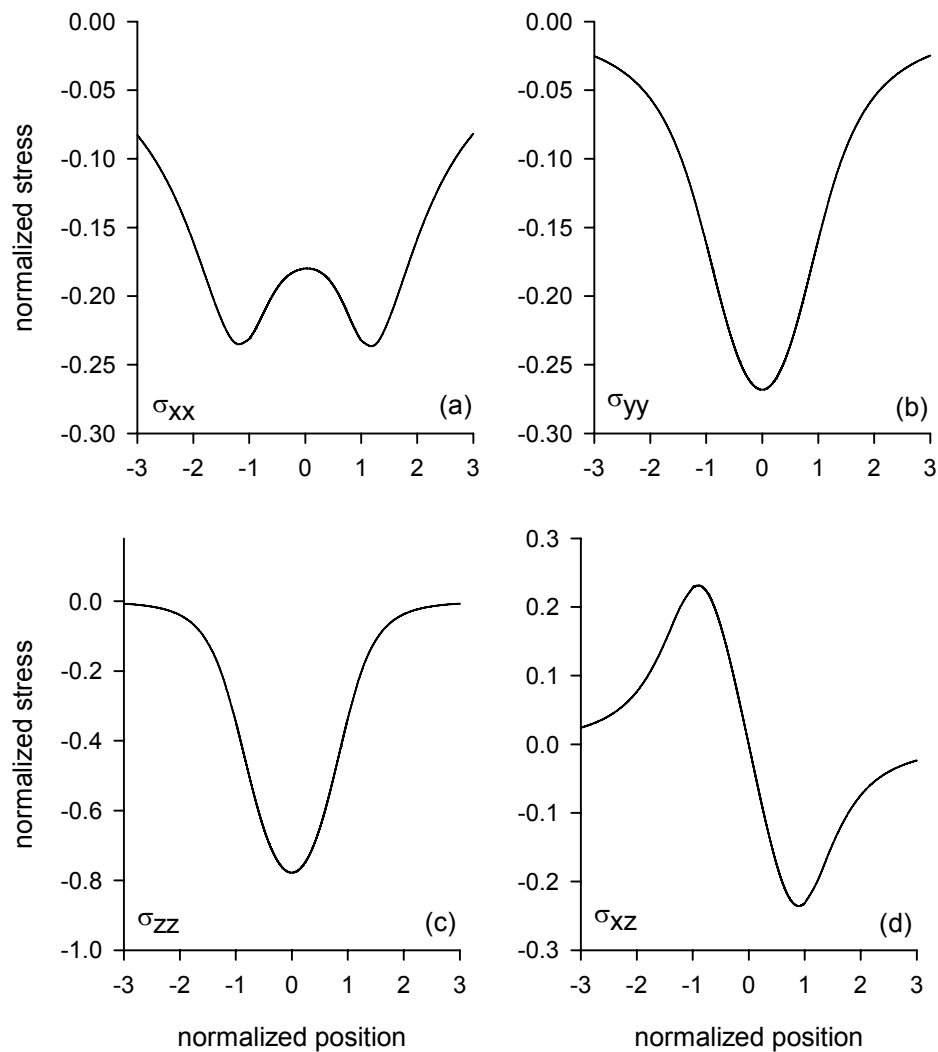


Figure 6.5.4.—Stress components as a function of normalized position for the normalized depth below surface of 0.8 units. Stress calculations based on plane strain response to applied loads, zero friction, and zero residual stresses. (a) Component in the direction of the gear tooth involute. (b) Component in the direction of the gear tooth face width. (c) Component in the direction of depth below the surface. (d) Shear component 'XZ'.

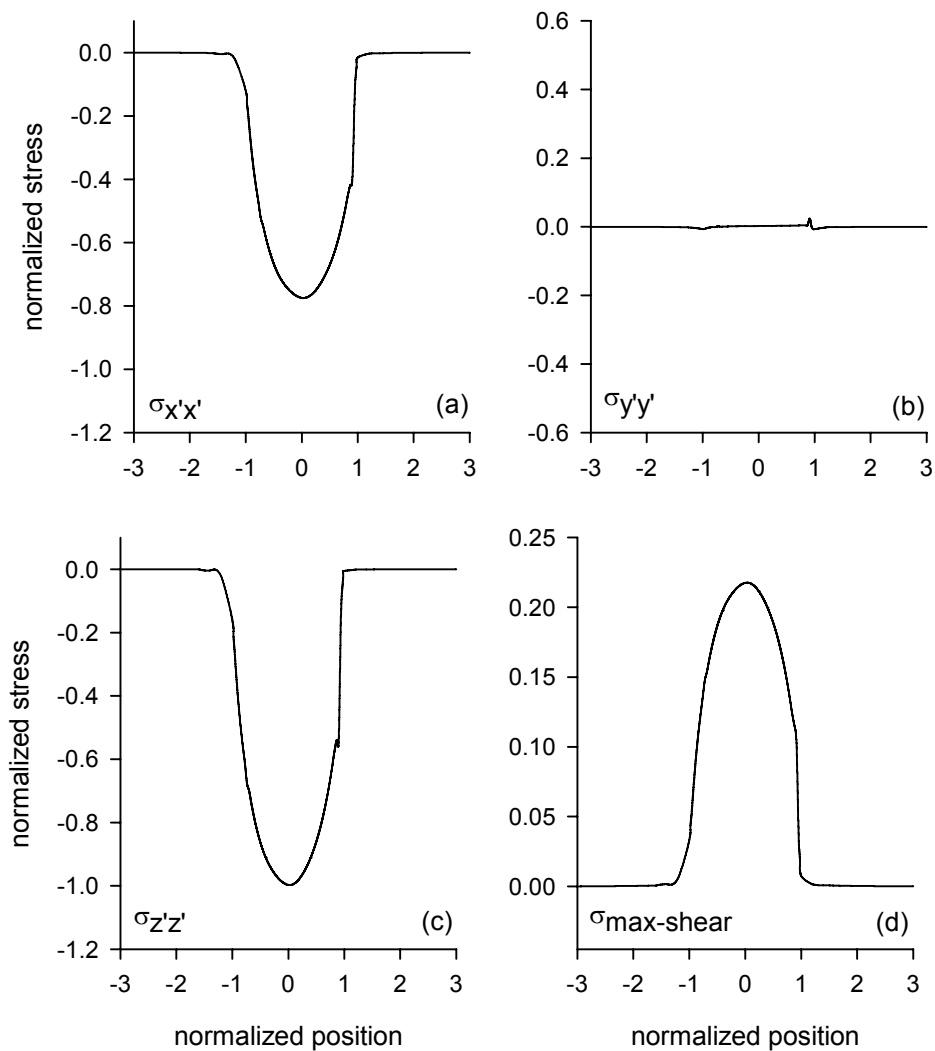


Figure 6.5.5.—Stress components as a function of normalized position for the normalized depth below surface of 0.004 units. Stress calculations based on plane strain response to applied loads, zero friction, and zero residual stresses. (a) Normal component $X'X'$, that is after coordinate system transformation (b) Normal component $Y'Y'$. (c) Normal component $Z'Z'$. (d) Shear component on plane of maximum shear.

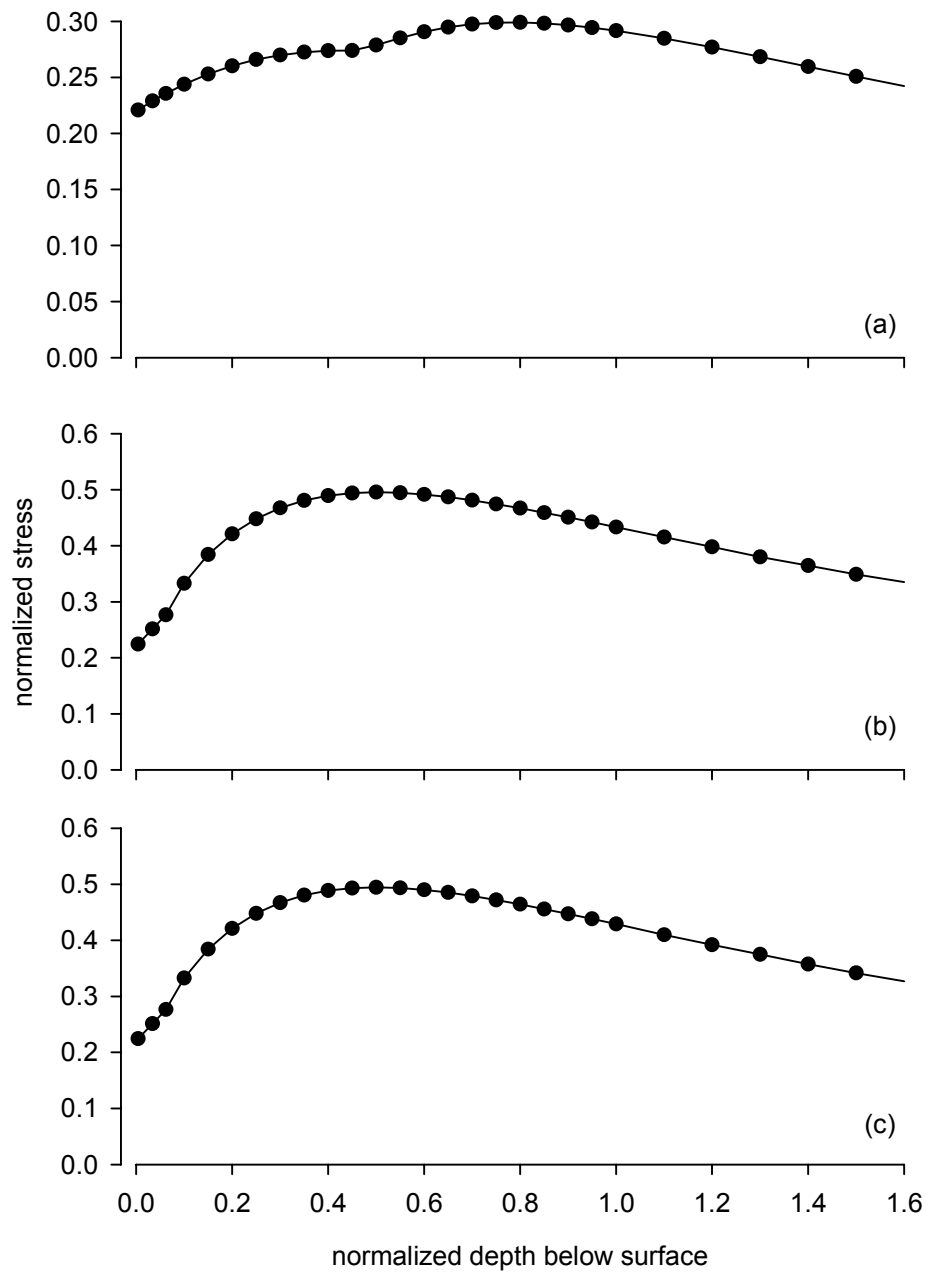


Figure 6.5.6.—Critical-plane stress-based indices of fatigue damage for baseline case of smooth lubricated surfaces, zero friction forces, and zero residual stresses. (a) Maximum shear value. (b) Maximum shear range. (c) Maximum values of the multiaxial fatigue parameter.

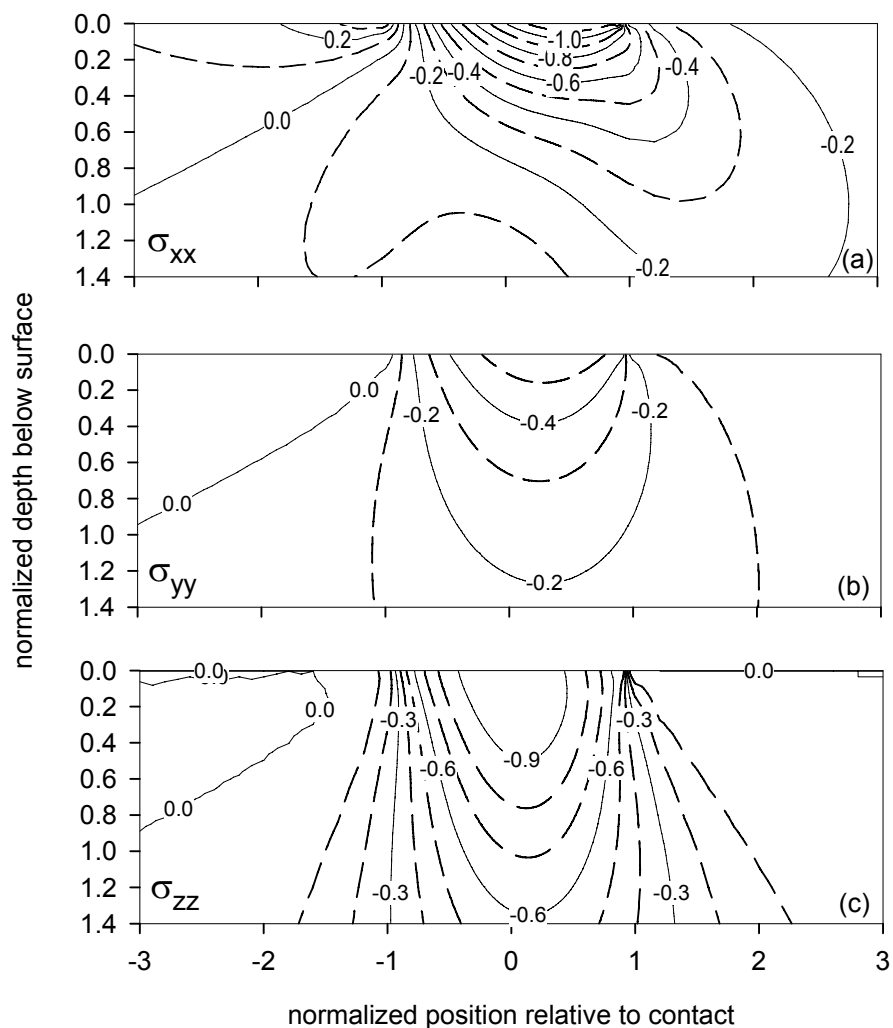


Figure 6.5.7.—Contour plots of normal stress components for case of smooth lubricated surfaces, friction coefficient of 1/3, and zero residual stresses. (a) Component in the direction of the gear tooth involute (and direction of friction). (b) Component in the direction of the gear tooth face width. (c) Component in the direction of depth below the surface.

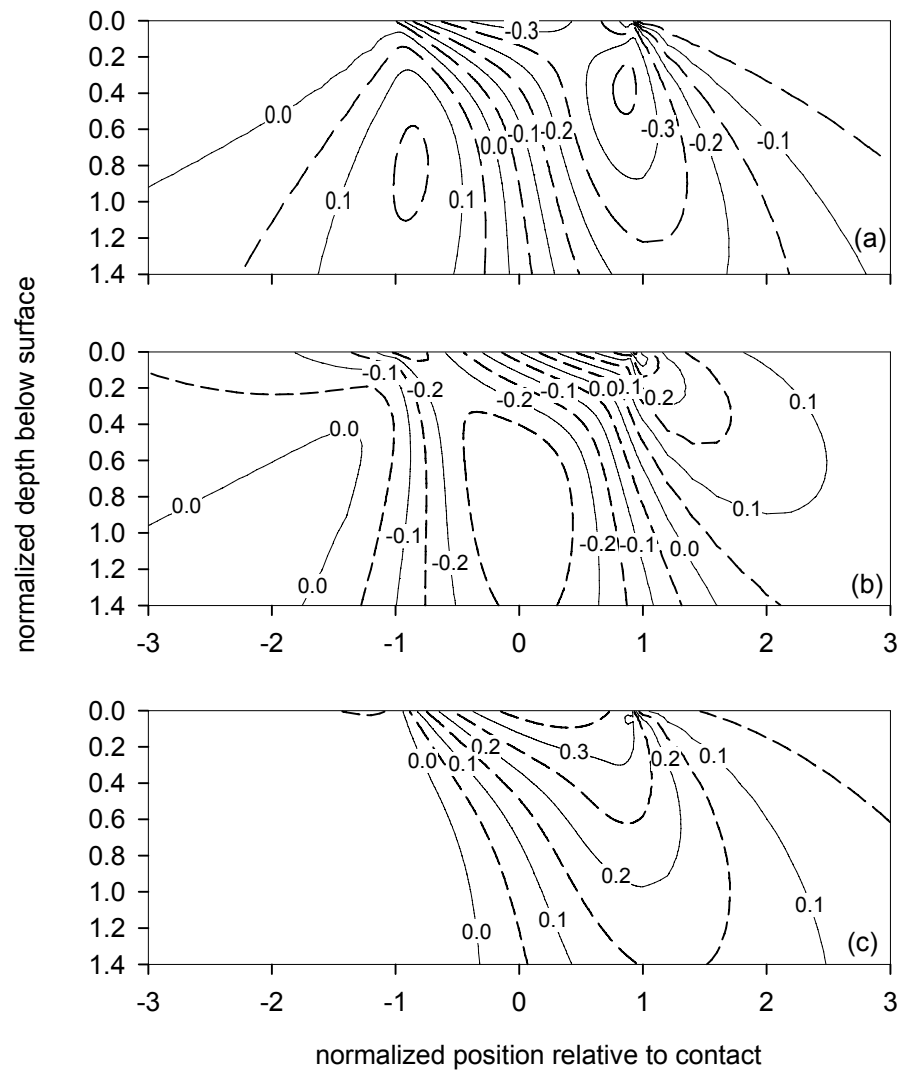


Figure 6.5.8.—Contour plots of shear stress components for case of smooth lubricated surfaces, friction coefficient of $1/3$, and zero residual stresses.
 (a) Shear component XZ of original coordinate system orientation.
 (b) Shear component on the plane of maximum shear range. (c) Shear component on the plane of maximum shear value.

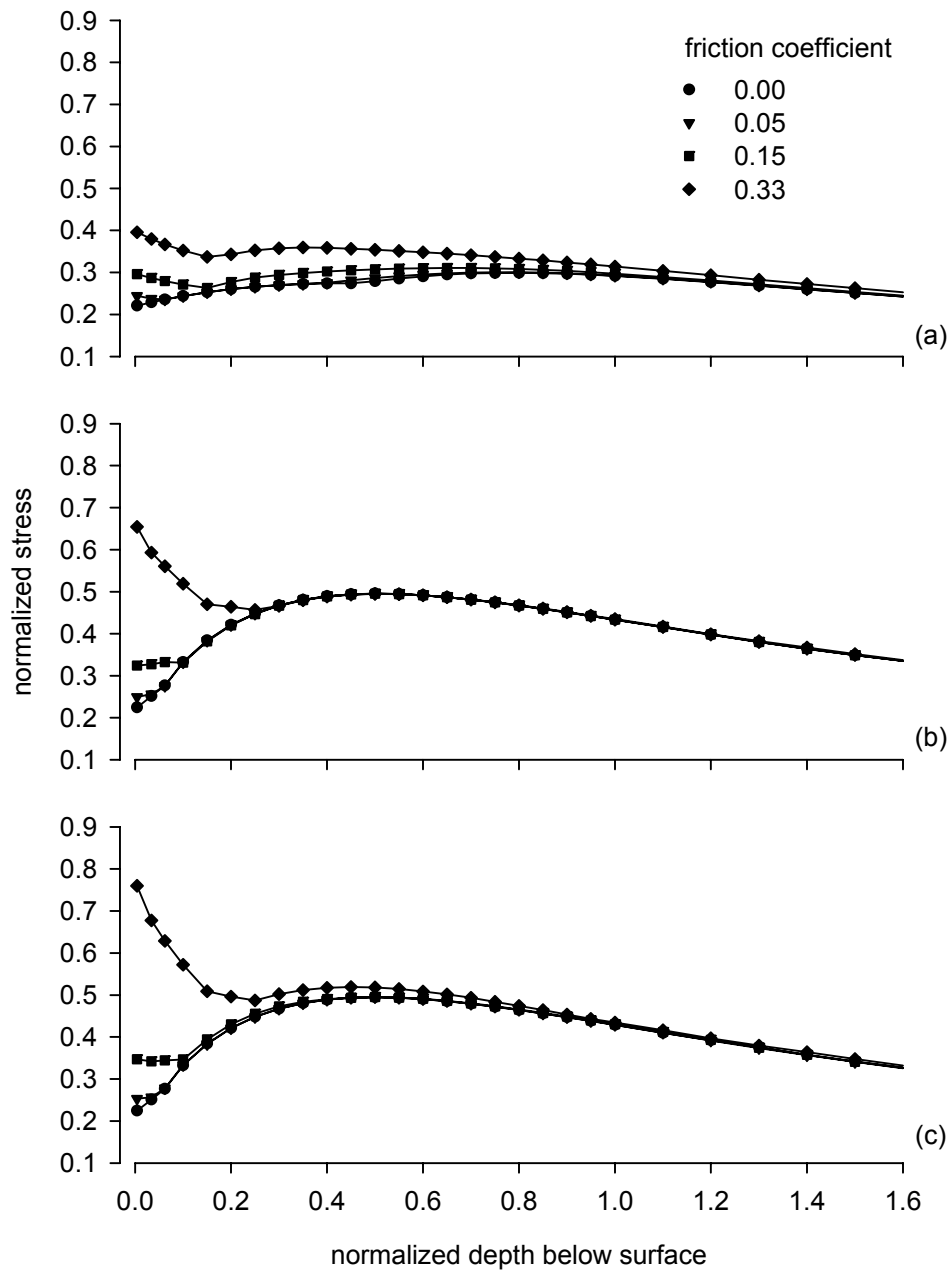


Figure 6.5.9.—Critical-plane stress-based indices of fatigue damage using four values of the friction coefficient and zero residual stresses. (a) Maximum shear values. (b) Maximum shear ranges. (c) Maximum values of the multiaxial fatigue parameter.

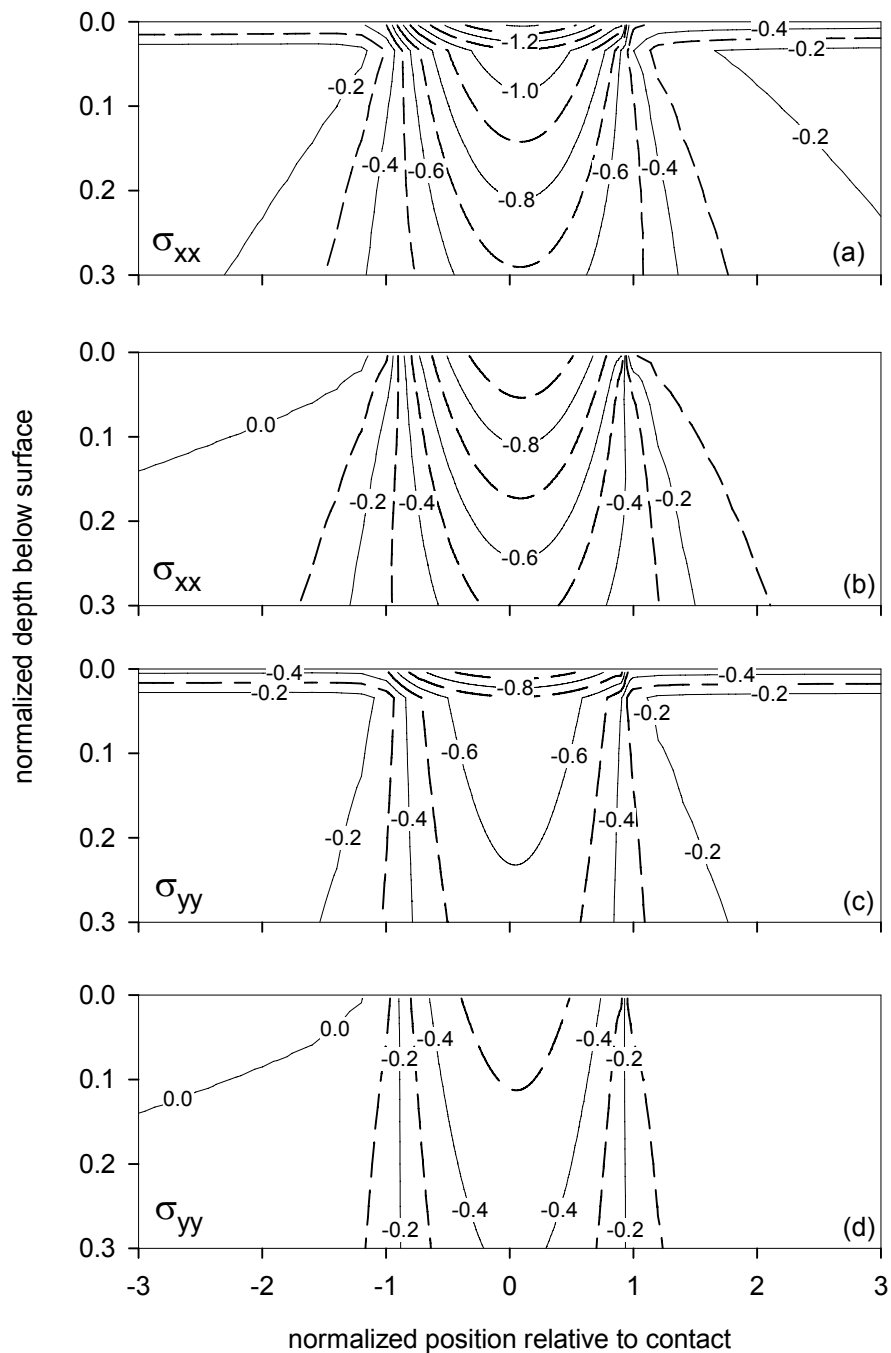


Figure 6.5.10.—Influence of residual stress field on two of the normal stress components. (a) Normal stress component in direction along involute, with residual stresses. (b) Normal stress component in direction along involute, without residual stresses. (c) Normal stress component in direction along tooth face, with residual stresses. (d) Normal stress component in direction along tooth face, without residual stresses.

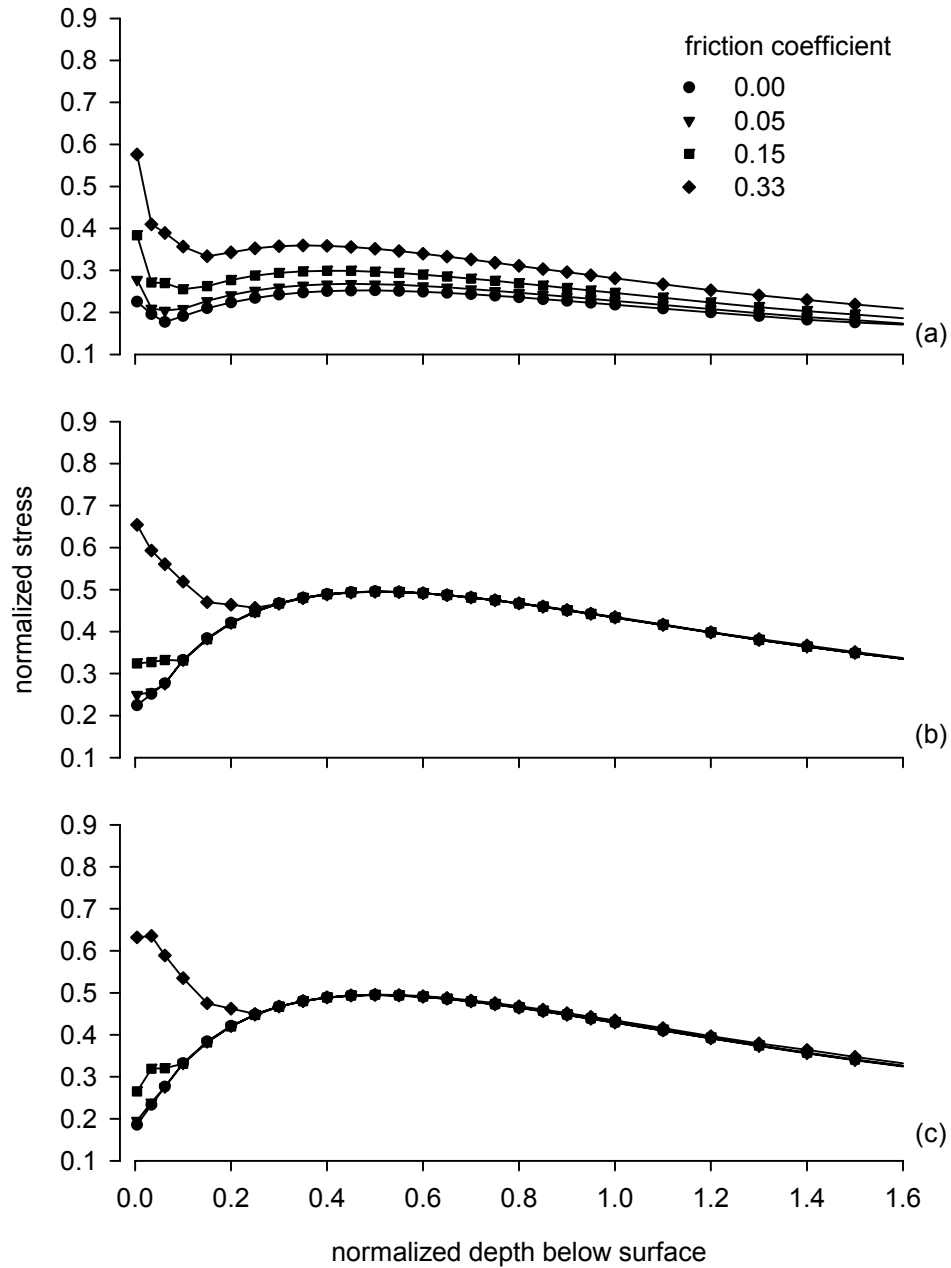


Figure 6.5.11.—Critical-plane stress-based indices of fatigue damage using four values of the friction coefficient and including the (superimposed) residual stress field and a plane strain response to applied loads. (a) Maximum shear values. (b) Maximum shear ranges. (c) Maximum values of the multiaxial fatigue parameter.

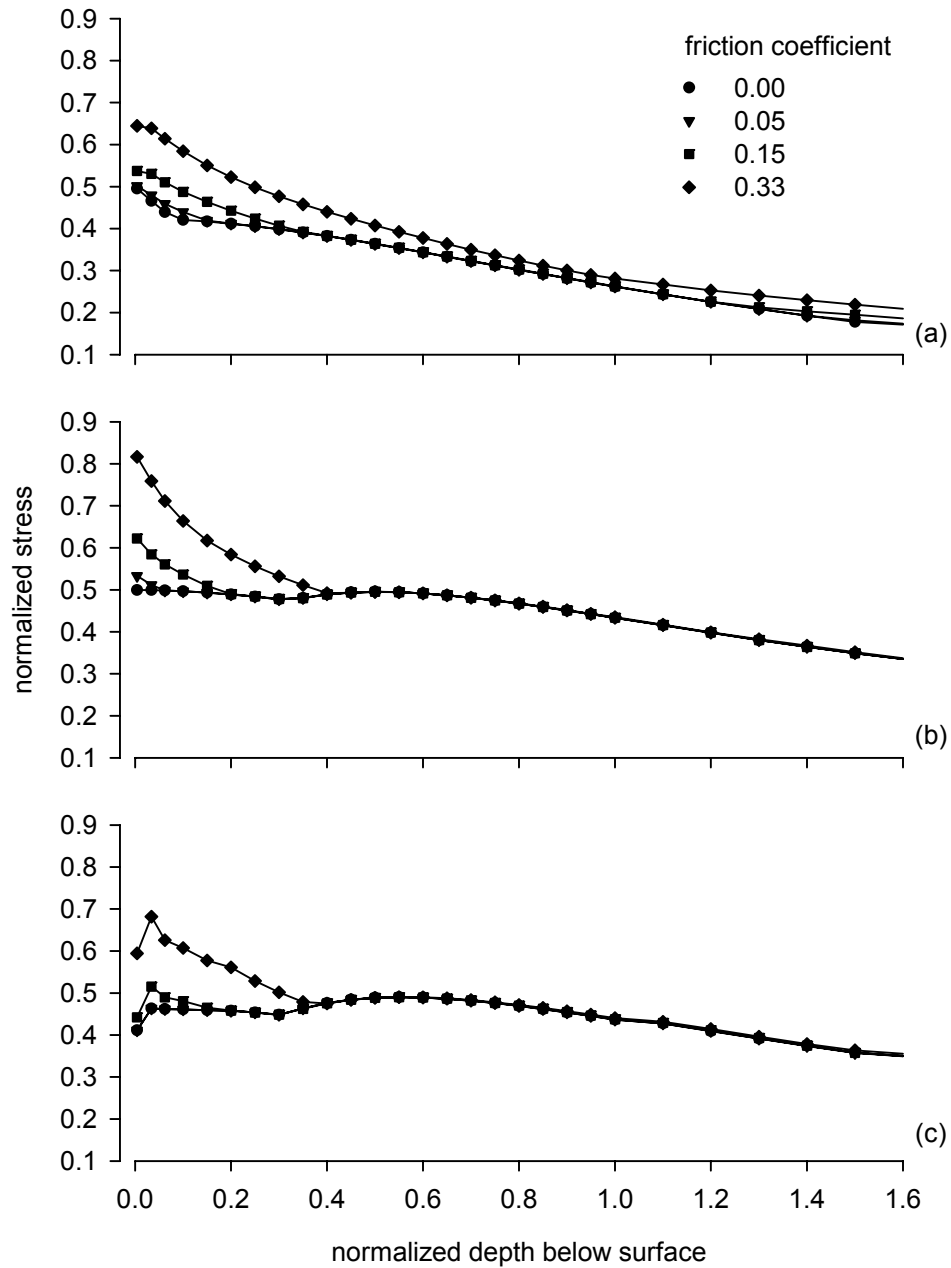


Figure 6.5.12.—Critical-plane stress-based indices of fatigue damage using four values of the friction coefficient and including the (superimposed) residual stress field and a plane stress response to applied loads. (a) Maximum shear values. (b) Maximum shear ranges. (c) Maximum values of the multiaxial fatigue parameter.

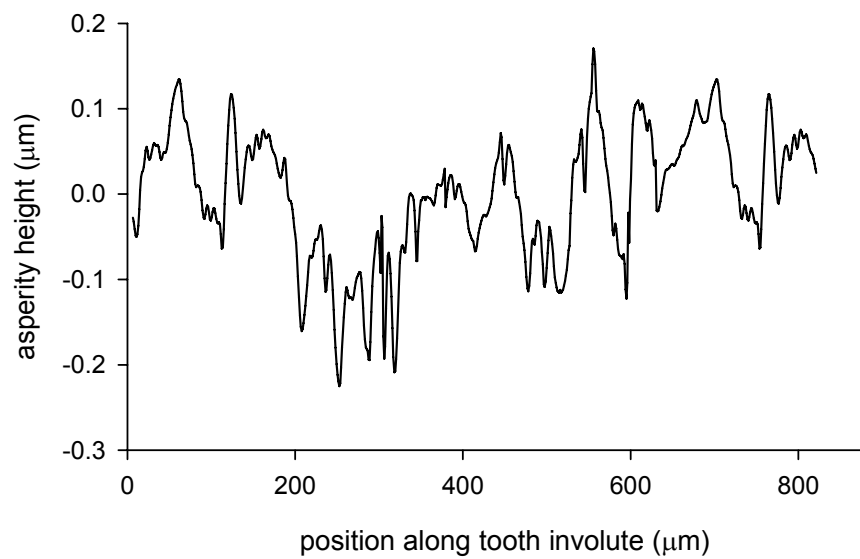


Figure 6.6.1.—Roughness profile used for rough surface elasto-hydrodynamic lubrication analysis. The vertical scale is approximately 1000X of the horizontal scale.

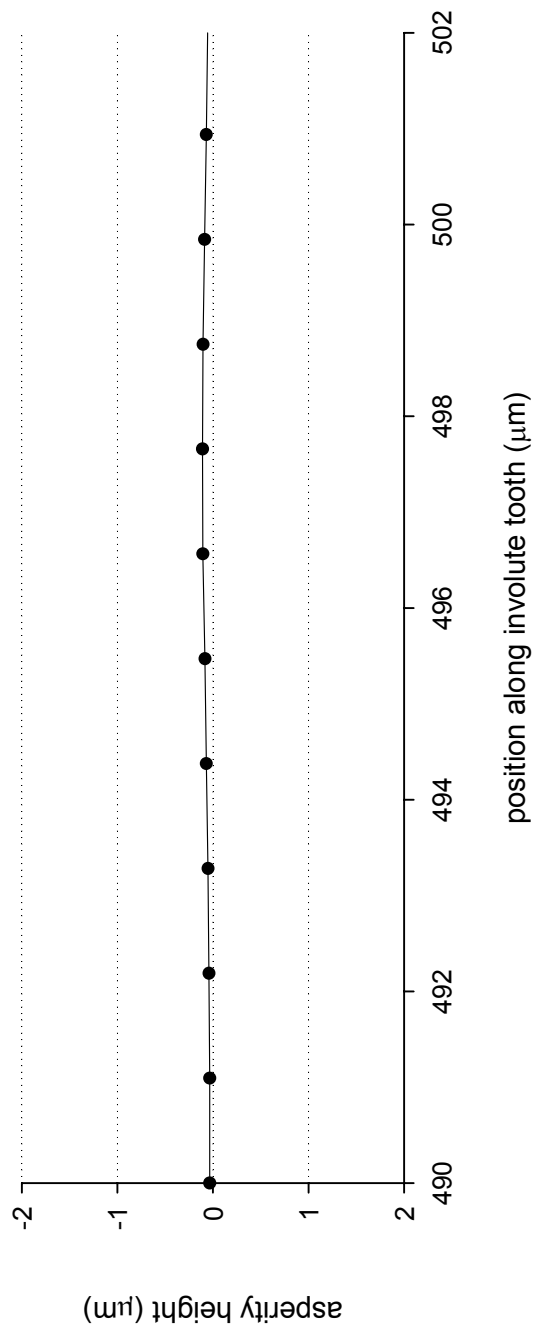


Figure 6.6.2.—Portion of roughness profile from figure 6.6.1 showing asperity geometry with true aspect ratio.

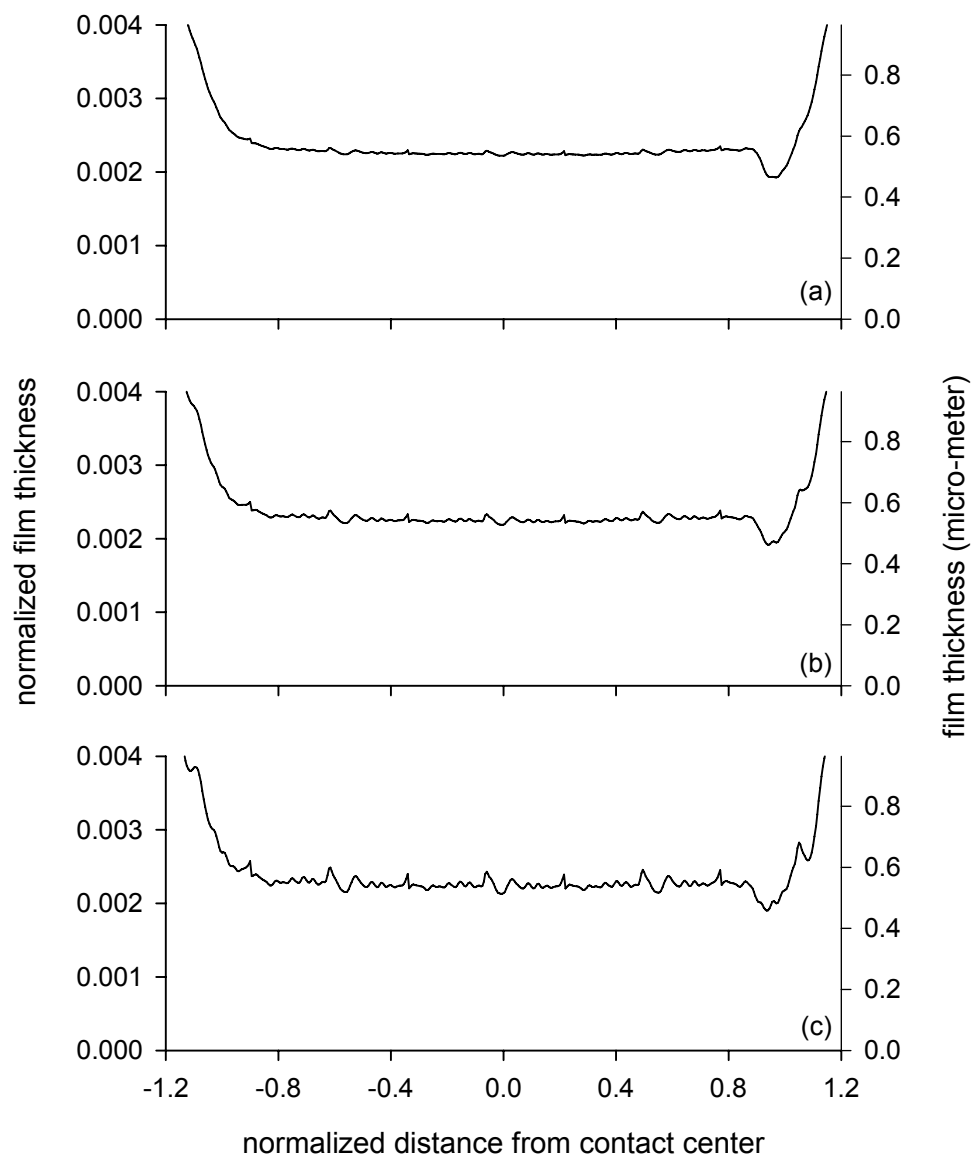


Figure 6.6.3.—Film thickness predictions by elastohydrodynamic analyses for three scaled versions of a composite superfinished roughness profile. (a) Solution for peak-to-valley roughness of 0.107 μm . (b) Solution for peak-to-valley roughness of 0.214 μm . (c) Solution for peak-to-valley roughness of 0.396 μm .

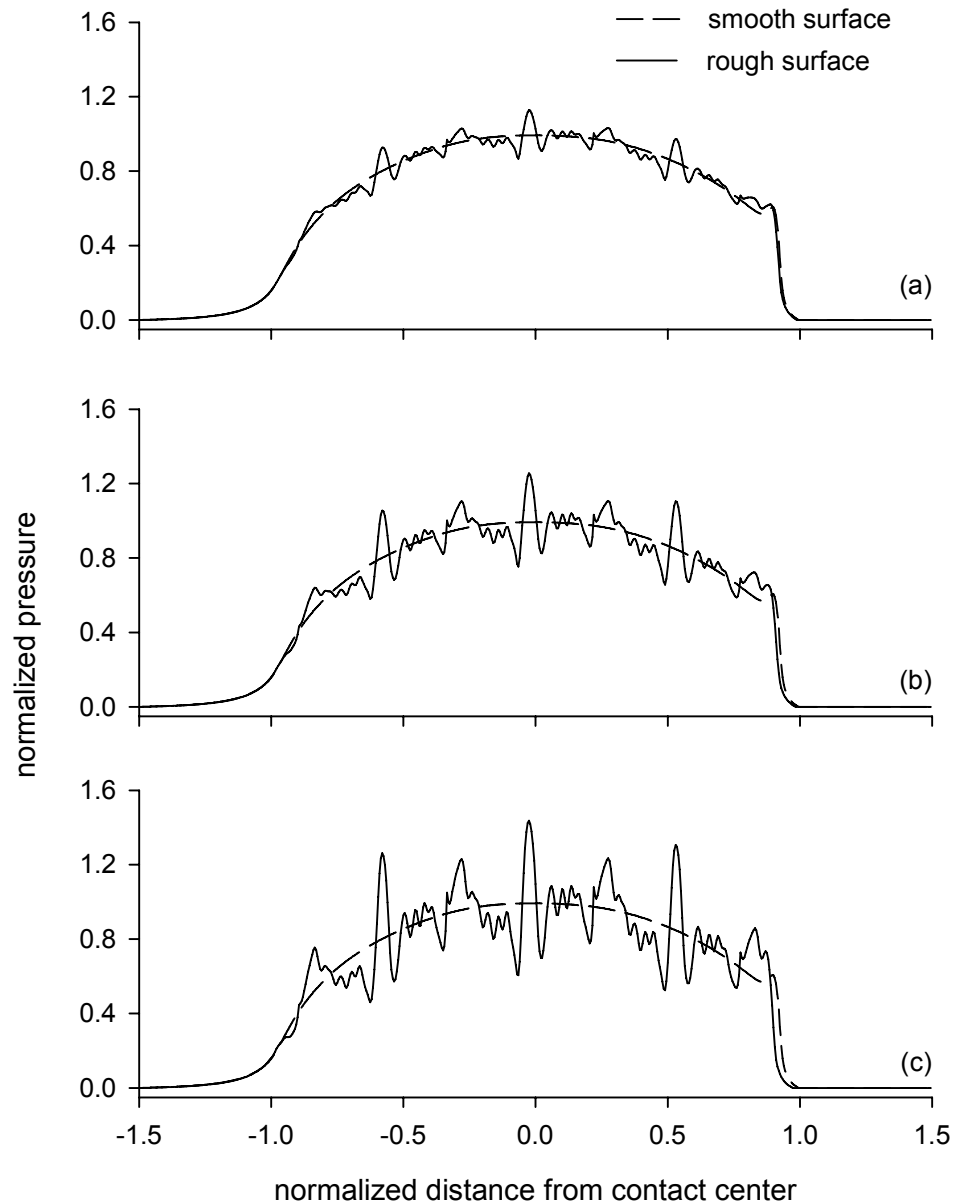


Figure 6.6.4.—Pressure distributions predicted by rough-surface elasto-hydrodynamic lubrication analysis for three scaled versions of a composite superfinished roughness profile. (a) Solution for peak-to-valley roughness of $0.107\ \mu\text{m}$. (b) Solution for peak-to-valley roughness of $0.214\ \mu\text{m}$. (c) Solution for peak-to-valley roughness of $0.396\ \mu\text{m}$.

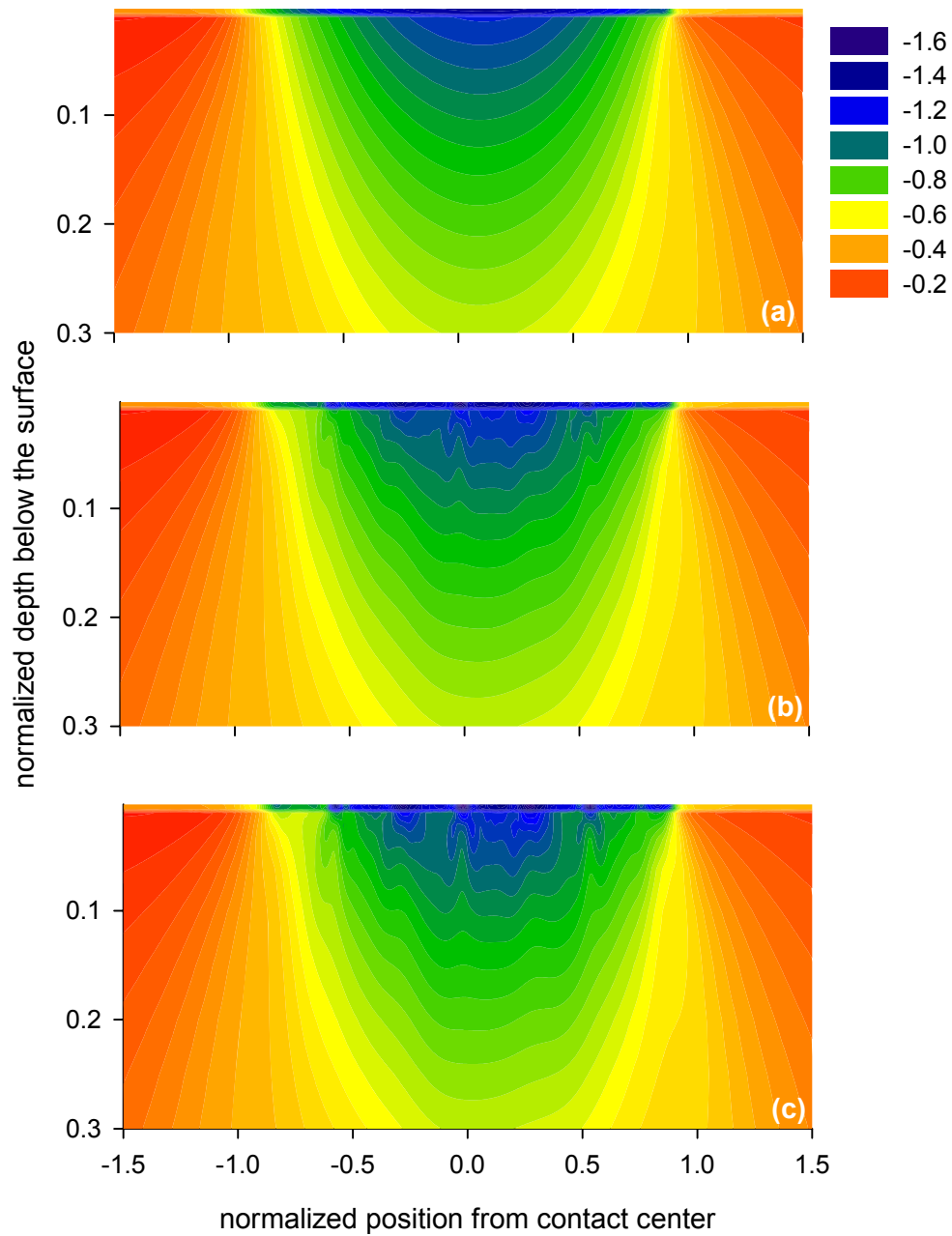


Figure 6.6.5.—Contours of normalized stresses for component σ_{xx} of the stress tensor as predicted by rough-surface elastohydrodynamic lubrication analysis for three scaled versions of a composite superfinished roughness profile. (a) Solution for ideally smooth surfaces. (b) Solution for peak-to-valley roughness of 0.214 μm . (c) Solution for peak-to-valley roughness of 0.396 μm .

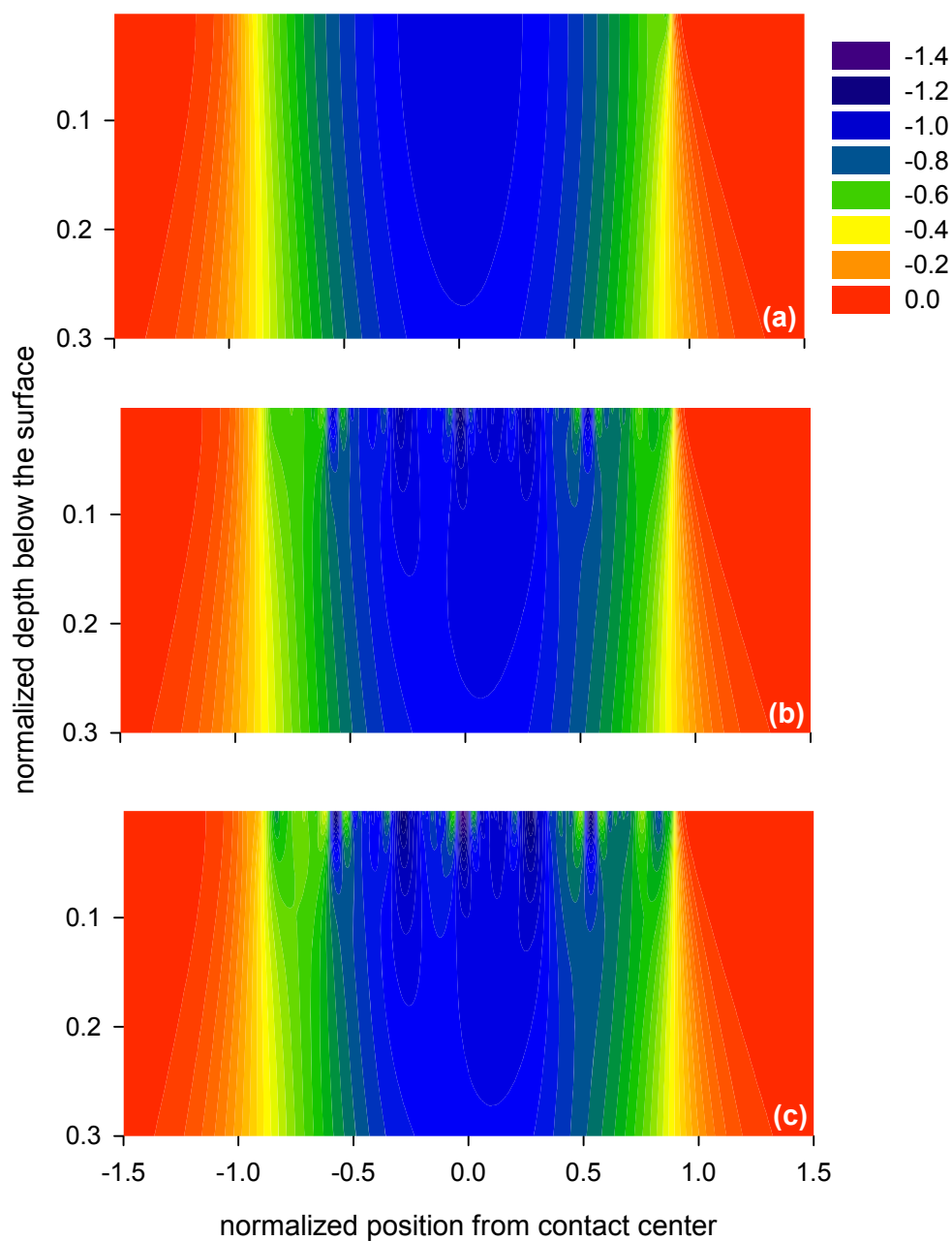


Figure 6.6.6.—Contours of normalized stresses for component σ_{zz} of the stress tensor as predicted by rough-surface elastohydrodynamic lubrication analysis for three scaled versions of a composite superfinished roughness profile. (a) Solution for ideally smooth surfaces. (b) Solution for peak-to-valley roughness of 0.214 μm . (c) Solution for peak-to-valley roughness of 0.396 μm .

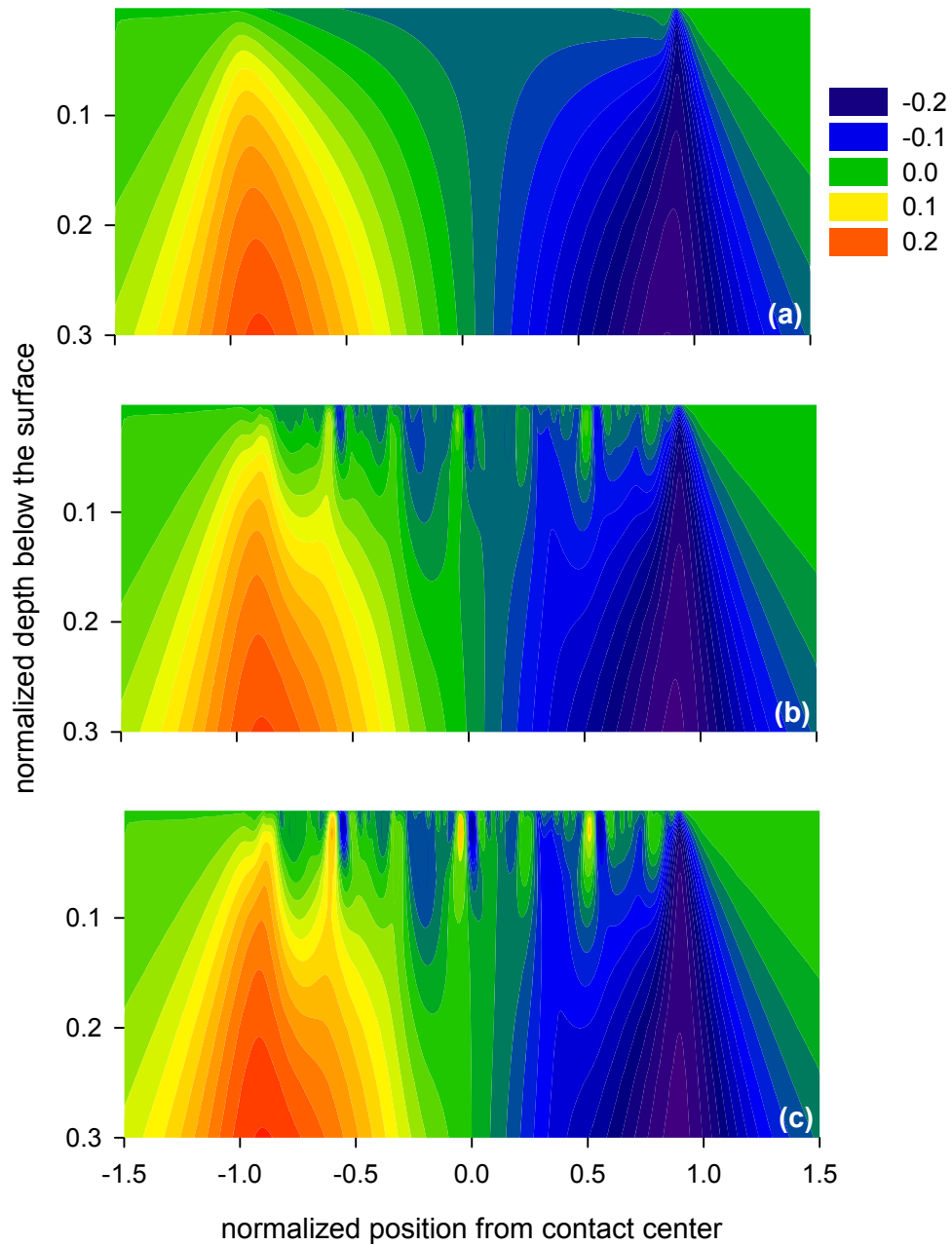


Figure 6.6.7.—Contours of normalized stresses for shear component σ_{xz} of the stress tensor as predicted by rough-surface elastohydrodynamic lubrication analysis for three scaled versions of a composite superfinished roughness profile. (a) Solution for ideally smooth surfaces. (b) Solution for peak-to-valley roughness of $0.214\text{ }\mu\text{m}$. (c) Solution for peak-to-valley roughness of $0.396\text{ }\mu\text{m}$.

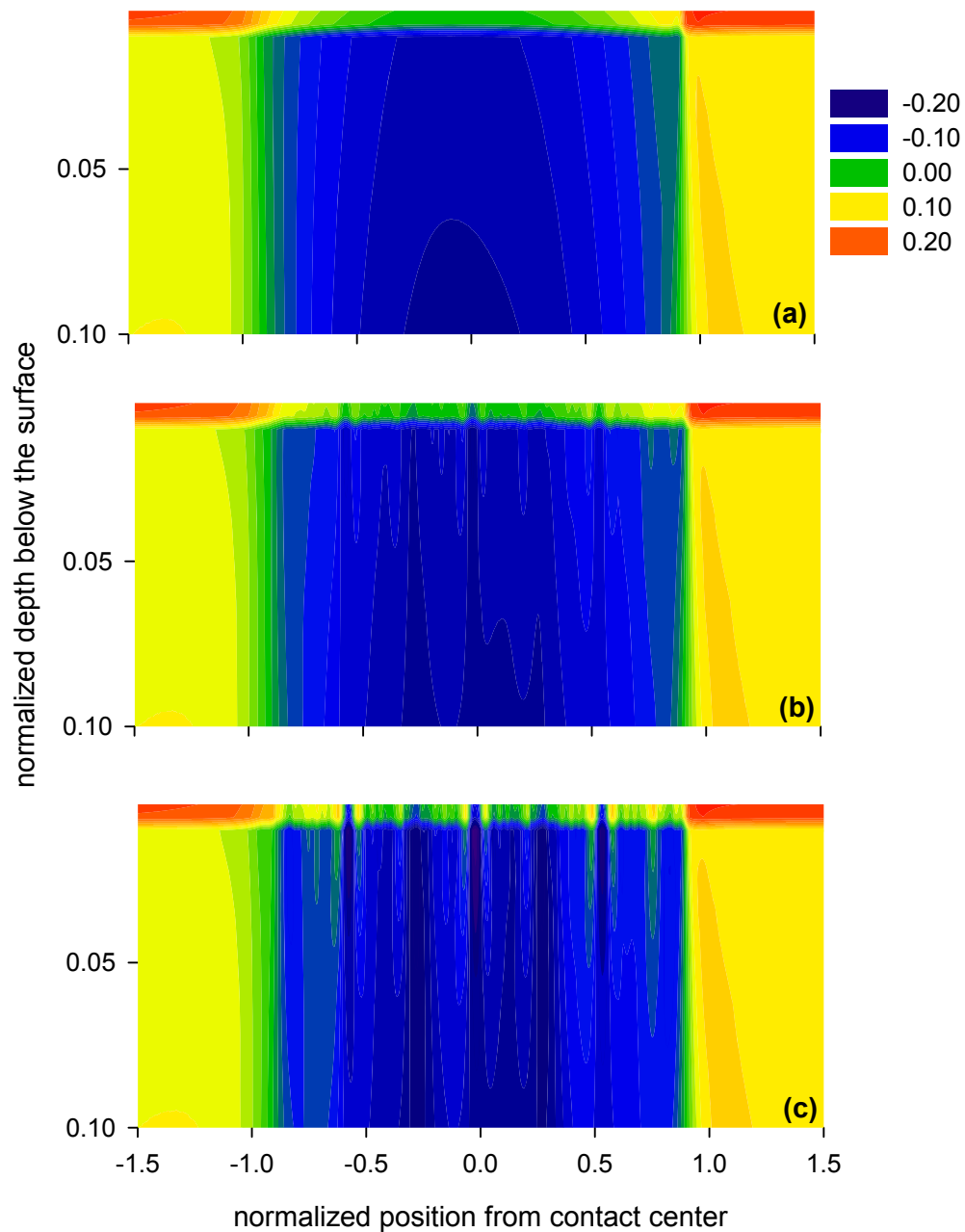


Figure 6.6.8.—Contours of normalized stresses for shear component, after transformation of the stress tensor by a 45 degree rotation about the ‘X’ axis, as predicted by rough-surface elastohydrodynamic lubrication analysis for three scaled versions of a composite superfinished roughness profile. (a) Solution for ideally smooth surfaces. (b) Solution for peak-to-valley roughness of 0.214 μm . (c) Solution for peak-to-valley roughness of 0.396 μm .

Chapter 7—Gear Tooth Surface Topography

7.1 Introduction

The experiments documented in chapter 4 showed conclusively that gears with differing as-manufactured surface topographies can have dramatically differing performance characteristics. Furthermore, it was observed that for the ground surfaces the surface topography changed with running. Sayles (ref. 7.1) considers that surface topography data combined with numerical contact mechanics have great potential for general use as production and design tools. To help provide insight about the relationship of the surface topography and the resulting performance as was demonstrated in the present work, inspections were completed for both as-manufactured and run-in surfaces. The instrument used for the inspections employs scanning white light interferometry technology (Ref 7.2).

7.2 Description of Inspection Method and Data Display

To accomplish surface topography inspections, gear teeth were selected and cut away from the body of the gear. The teeth that were inspected were selected as being representative samples. The gear tooth samples were cleaned before submitting for inspection. The inspections were accomplished using a lens and resolution setting that provided an overall system magnification of 200X. The field of view was 0.700 mm by 0.525 mm. The lateral discrete sampling spacing was 1.1 μm . The vertical resolution of the inspection machine was 0.1 nm

The arrangement of the data display from the mapping interferometric microscope is provided in figure 7.2.1. The intensity map is located toward the bottom—right corner of the display. The intensity map is an optical image used by the operator to focus the instrument and to orient the surface to be normal to the optical measuring path. When the data acquisition is completed, the processed data is displayed as a surface plot, an oblique plot, and a profile plot. For all measurements provided in this work, the instrument software removed a cylindrical form from the data displayed in these three plots. Removing the cylindrical form allows for a more detailed display of waviness and roughness features.

The surface plot is located toward the upper—left quadrant of the display (fig. 7.2.1). The surface plot is a color-coded plot of surface heights. Auto-scaling was used to scale the color coding. The scaling is noted just to the left of the three-dimensional surface plots. Since auto-scaling was employed, one needs to use care in comparing one surface inspection to another. Surface statistics are displayed just below the surface plot. The PV value indicates the maximum peak-valley separation, “r.m.s.” indicates the surface root-mean-squared value from the mean plane, and Ra indicates the surface roughness average from the mean plane. No additional processing of the data was accomplished, beyond the previously mentioned removal of the cylindrical form, for all work presented here. Therefore, the data includes “spurious” readings due to the limitations of the inspection technique. Since the “PV” peak-to-valley separation distance brackets any spurious data points, this statistic might not be representative of the surface. The root-mean-squared and roughness average statistics will also be affected by such spurious readings, but since the surface data contains relatively few of these points the influence of these spurious readings on the root-mean-squared and roughness average values can be considered minor.

Also located on the surface plot is a line bound by two triangles (fig. 7.2.1). The operator is able to manipulate the display to place this “profile line” at a desired location and orientation. The profile line provides a reference showing the data that is displayed on the profile plot.

The profile plot is located toward the lower-left quadrant of the display (fig. 7.2.1). The location of the profile that is displayed is defined by the profile line on the surface map. Auto-scaling was employed, and so one needs to use care in comparing one profile inspection to another. Located just below the profile plots are a listing of statistics of the profile data. As mentioned in the discussion of the surface plot, the surface measurement data might include spurious data. Since the PV peak-to-valley separation distance brackets any spurious data points this statistic might not be representative of the profile. The root-mean-squared and roughness average values for the displayed profile are also listed. The root-mean-

squared and roughness average statistics will also be affected by any spurious data, but since the profile data contains relatively few of these spurious data points the influence of these data can be considered as minor.

The oblique plot is located toward the upper-right quadrant of the data display. The data for the oblique plot is the same as the data for the surface plot but displayed in another manner. The oblique plot provides for a visual image of the surface texture. For the case of ground surfaces, the orientation of the grinding furrows are clearly displayed.

The orientation of the gear tooth relative to the measuring machine's coordinate system was as denoted on the bottom of figure 7.2.1. In all cases, the positive x-direction of the measuring machine's coordinate system was approximately aligned to be along the gear tooth involute profile direction, and the positive x-direction was oriented from the gear tooth root toward the gear tooth tip.

7.3 Inspection Results for the As-Manufactured Condition—Ground Gears

In this project both as-manufactured and run-in areas of gear tooth surfaces were inspected. For the case of superfinished gear teeth, one could not make a distinction between the as-manufactured and run-in areas on the basis on the inspection data. Therefore, for the superfinished surfaces only inspections for run-in areas are included in this document (see section 7.5). For the case of ground surfaces, the surface texture changed dramatically due to the running-in of the surfaces. Data from the un-run areas of two gears are provided and discussed in the paragraphs to immediately follow.

The surface inspection data display for ground gear no. 1 is provided in figure 7.3.1. The location for the inspection was approximately the middle of the un-run portion of the face-width and at the pitch point location on the involute profile. The orientation of the grinding furrows are at an angle of about 30 degrees relative to the direction across the face width (that is relative to the y-axis of the measuring machine's coordinate system). The roughness average for the selected profile is $0.47\text{ }\mu\text{m}$ (18 $\mu\text{in.}$).

The surface inspection data display for ground gear no. 2, area no. 1 is provided in figure 7.3.2. The location for the inspection was approximately the middle of the face-width and at the pitch point location on the involute profile (the same as for gear no. 1 described in the preceding paragraph). As was the case for gear no. 1, the orientation of the grinding furrows is at an angle of about 30 degrees relative to the direction across the face width. The roughness average for the selected profile is $0.57\text{ }\mu\text{m}$ (22 $\mu\text{in.}$).

The surface inspection data display for ground gear no. 2, area no. 2 is provided in figure 7.3.3. The location for the inspection was approximately the middle of the face-width and somewhat above the pitch point location, between the tooth tip and the high-point of single-tooth-contact. The orientation of the grinding furrows is at an angle of about 45 degrees relative to the direction across the face width. The roughness average for the selected profile is $0.48\text{ }\mu\text{m}$ (19 $\mu\text{in.}$).

As is evident from these inspections, the orientation of the grinding furrows depends on the location of the inspected area (both the position across the face and position along the involute profile direction). This changing orientation is due to a circular-shaped motion and small size of the grinding wheel. The grinding furrow patterns were visible on the un-run portions of the gear teeth and were similar to that displayed in figure 4.3.1.

7.4 Inspection Results for the Run-In Condition—Ground Gears

Surface inspections were completed for the run-in surface of one of the ground gears used for the experiments described in chapter 4. The surface that was inspected had both unrun and run-in portions of the face due to the face off-set testing method. The run-in portion of the tooth had endured 57 million engagement cycles at a load intensity of 580 N/mm (3300 lb/in.). The testing of this gear had been suspended due to fatigue failure of another tooth on the gear, but the inspected surface had no visible fatigue spalls, pits, nor any other fatigue failure features of size to be visible. This gear tooth was selected as having a typical visual appearance for the run-in ground gears.

The surface inspection data display for the run-in ground gear, area no. 1 is provided in figure 7.4.1. The location for this inspection was approximately the middle of the running track across the facewidth and somewhat above the pitch point location, between the tooth tip and the high-point of single-tooth-contact. The surface heights plot and the oblique plot show that there has been a significant change in the surface topography due to the sliding, loaded contact. The grinding furrows that were clearly revealed by inspection of the un-run surface are no longer evident. The profile plot, oriented along the tooth involute profile, shows evidence of some relatively sharp, deep valley features that may be the remnants of the grinding marks. The magnitude of the roughness, as characterized by either of the root-mean-squared or roughness average, has been reduced relative to the ground surface. There is evidence of significant wear and/or plastic deformation of the asperity tips. The slopes of the asperities have been considered by some researchers as an especially important surface characteristic, with sharper asperities being correlated to shorter fatigue lives (refs. 7.3 and 7.4). The plastic deformation and resulting redistribution of residual stresses in the neighborhood of surface asperities has also been studied and proposed as a controlling factor for surface fatigue (ref. 7.5).

The surface inspection data display for the run-in ground gear, area no. 2 is provided in figure 7.4.2. The location for this inspection was approximately the middle of the running track across the facewidth and approximately at the pitch point location along the involute profile. The surface heights plot and the oblique plot shows significant change in the surface topography as was seen in the inspection toward the tooth tip. Toward the right edge of the data plots is evident a feature resembling a raised ridge. This feature has a wavelength of about 0.2 mm (0.008 in.), a long wavelength relative to roughness features. It is speculated that this area may be the location of pure rolling along the involute. Some models for gear wear suggest zero wear for a pure rolling condition, and so the ridge may indicate low wear rates at this location. However, we also note that the top of the ridge does not show evidence of grinding furrows, and so even at this location of speculated low wear there has been some change of the surface topography.

The surface inspection data display for the run-in ground gear, area no. 3 is provided in figure 7.4.3. The location for this inspection was approximately the middle of the running track across the facewidth and below the pitch point location along the involute profile near the low-point of single-tooth-contact. The surface heights plot and the oblique plot shows significant change in the surface topography as was seen in the other inspection toward the tip and near the pitch point. The small-scale topography features and surface statistics are similar for all of the inspected run-in locations along the involute profile.

Perhaps the most striking feature of the just described three inspections was the raised ridge seen in figure 7.4.2. Therefore, additional inspections were completed at differing position across the face width but at the same location along the involute near the pitch point. The inspection included one inspection of the un-run portion of the gear tooth. The oblique plots were gathered together and are displayed in figure 7.4.4. The “raised ridge” feature is evident in all of the inspections of the run-in surface, but it is not evident on the un-run surface, figure 7.4.4(a).

The inspections of the run-in surfaces show significant modification of the small-scale surface geometry features at all locations on the tooth face. To display the differences of the un-run and the run-in surfaces, inspections were conducted at the edges of the running track. Two such inspections are provided in figures 7.4.5 and 7.4.6. There exists at the edge of the running track an interesting feature resembling a channel and an adjacent raised ridge oriented along the involute profile direction. Note that on these two plots, the profile line was oriented to be at right angles to the channel-like feature. The grinding furrows just adjacent to the edge of the running track are evident. It is not clear whether the channel-like feature is the result of wear, plastic flow, or some combination of these two phenomena. Edge effects for line contacts has been studied, and increased pressures at edges of line contacts has been predicted (ref. 7.6) as would be consistent with the present observations. This channel-like feature might also be related to thinning of films at the edges of the line contacts due to pressure gradients and side leakage from valley features. Such action has been speculated by Patching, et al. (ref. 7.7) as an explanation of scuffing (or scoring) phenomena.

From all of these inspections, it is evident that for the experimental conditions of chapter 4 used to conduct fatigue testing, there exists a combination of wear and fatigue actions. Most often these are

considered separate modes of (potential) gear failure, but it is also probable that wear influences fatigue (ref. 7.8). Plastic deformation also likely occurs (at least to some degree) in many practical gearing applications, and this phenomena also likely influences the fatigue mechanisms and fatigue life (refs. 7.5, 7.9, to 7.11). A comprehensive model for gear durability considering all of wear, plastic flow, and fatigue remains the subject of further research efforts by the gearing and tribology technical communities.

7.5 Inspection Results For the Run-In Condition—Superfinished Gears

Surface inspections were completed for the run-in surface of one of the superfinished gears used for the experiments described in chapter 4. The surface that was inspected had both portions of the tooth face tested and a relatively narrow track in the middle of the tooth that represents the un-run surface. This gear had endured 300 million engagement cycles at a load intensity of 580 N/mm (3300 lb/in.). The testing of this gear had been suspended due to a preselected fatigue test suspension time. There were no visible fatigue spalls, pits, nor any other fatigue failure features of size to be visible on any teeth of this gear. The gear tooth was selected as having a typical visual appearance for the run-in superfinished gears.

The surface inspection data display for the run-in superfinished gear, area no. 1 is provided in figure 7.5.1. The location for this inspection was approximately the middle of the running track across the facewidth and somewhat above the pitch point location, between the tooth tip and the high-point of single-tooth-contact. One should note that the autoscaling feature provides for smaller overall scales of the surface height features as compared to the inspections for the ground, run-in surfaces. The surface heights plot and the oblique plot show that there has little change of the surface topography due to the sliding, loaded contact. The tested superfinished surfaces appear very much like un-run surfaces. A trace representing a deep grinding mark feature that was not removed by the superfinishing process is evident toward the left portion of the surface plots. The grinding mark was not removed by wear with running even after 300 million cycles. This would imply that for the superfinished surfaces and the experimental conditions for fatigue testing used in this work, the amount of wear of the gear tooth surfaces was negligible.

The surface inspection data display for the run-in superfinished gear, area no. 2 is provided in figure 7.5.2. The location for this inspection was approximately the middle of the running track across the facewidth and approximately at the pitch point location along the involute profile. The surface features are similar to those of inspection area no. 1 for this tooth. The “raised ridge” feature that is speculated to be a feature of the pitch point of the ground and run-in gears (fig. 7.4.4) could not be found on the superfinished and run-in gear.

The surface inspection data display for the run-in superfinished gear, area no. 3 is provided in figure 7.5.3. The location for this inspection was approximately the middle of the running track across the facewidth and below the pitch point location along the involute profile near the low-point of single-tooth-contact. As was true for areas no. 1 and no. 2 of the superfinished gear, there is no evidence of significant changes to the surface topography due to running-in, and amounts of wear appear to be negligible.

For the case of run-in ground gears, a channel-like feature was observed at the edges of the running track. The edges of the running tracks for the superfinished gears were also inspected, but such a channel like feature could not be found (fig. 7.5.4). These observations are consistent with those of Patching, et al. They observed in their experiments for scuffing (scoring) failures that ground surfaces failed at the edges of the running track where, they speculated, side leakage of lubricant allowed for localized thinning of the film and opportunities for additional metal-to-metal contact. On the other hand, in their testing of scuffing of superfinished surfaces with mirror-like quality, scuffing occurred at the middle of the contact. They speculated that the lack of deep and numerous valley features helped to maintain the lubricant at the edges contacts for the superfinished surfaces. Some attempts have been made to model these speculations using fully-coupled micro-elastohydrodynamic analysis (ref. 7.12).

7.6 Surface Inspections—Summary

The surface inspections presented in this chapter showed that for the experimental conditions used in the present work for fatigue testing, the surface topography of the ground gears changed dramatically with running-in while the changes in surface topography for the superfinished gears was negligible. These inspections also perhaps raise more questions than answers concerning the interaction of wear, plastic deformation, and fatigue phenomena. The inspections and discussions presented in this chapter help to fully document the present work and, perhaps, will provide qualitative guidance for future theoretical and modeling treatments of gear contacts.

References

- 7.1. Sayles, R., "How Two- And Three-Dimensional Surface Metrology Data Are Being Used To Improve The Tribological Performance And Life Of Some Common Machine Elements," *Tribology International* **34** (5), (2001).
- 7.2. "NewView 5000 Operating Manual," OMP-0423G, Zygo Corp., Middlefield, CT., (2001). Available on-line, http://support.zygo.com/home/site_dir.cgi, accessed April 24, 2002.
- 7.3. Tallian, T., "The Influence of Asperity Statistics on Surface Distress and Spalling Life of Hertzian Contacts," *Tribology Transactions*, **36** [1], (1993).
- 7.4. Tallian, T., "The Evolution of Rolling Bearing Life Models," *Society of Tribologists and Lubrication Engineers*, Park Ridge, IL, (2000).
- 7.5. Olver, A., Spikes, H., Bower, A., and Johnson, K., "The Residual Stress Distribution in a Plastically Deformed Model Asperity," *Wear*, **107**, (1986).
- 7.6. Hartnett M., Kannel, J. "Contact Stresses Between Elastic Cylinders - A Comprehensive Theoretical And Experimental Approach," *Trans. of the ASME, Journal Of Lubrication Technology*, **103** [1], (1981).
- 7.7. Patching M., Kweh C., Evans H., Snidle R., "Conditions For Scuffing Failure Of Ground And Superfinished Steel Disks At High Sliding Speeds Using A Gas-Turbine Engine Oil," *Trans. of the ASME, Journal Of Tribology*, **117** [3], (1995).
- 7.8. Fan, H., Keer, L., Cheng, W., Cheng H., "Competition Between Fatigue Crack Propagation and Wear," *Journal of Tribology*, **115**, (1993).
- 7.9. Kapoor, A., Johnson, K., "Effect of Changes in Contact Geometry on Shakedown of Surfaces in Rolling/Sliding Contact," *International Journal of Mechanical Science*, **34** [3], (1992).
- 7.10. Johnson, K., Shercliff, H., "Shakedown of 2-Dimensional Asperities in Sliding Contact," *International Journal of Mechanical Science*, **34** [5], 1992
- 7.11. Kapoor, A., Morales-Espejel, G., Olver, A., "A Shakedown Analysis of Simple Spur Gears," *Tribology Transactions*, **45** [1], (2002).
- 7.12. Elcoate, C., Evans, H., Hughes, T., "Fully Coupled Elastohydrodynamic Solution Techniques for the Analysis of Real Rough Line Contacts Using Finite Element and Finite Difference Models," *Elastohydrodynamics '96 – Proceedings of the 23rd Leeds-Lyon Symposium on Tribology*, D. Dowson, ed., Elsevier, (1997).

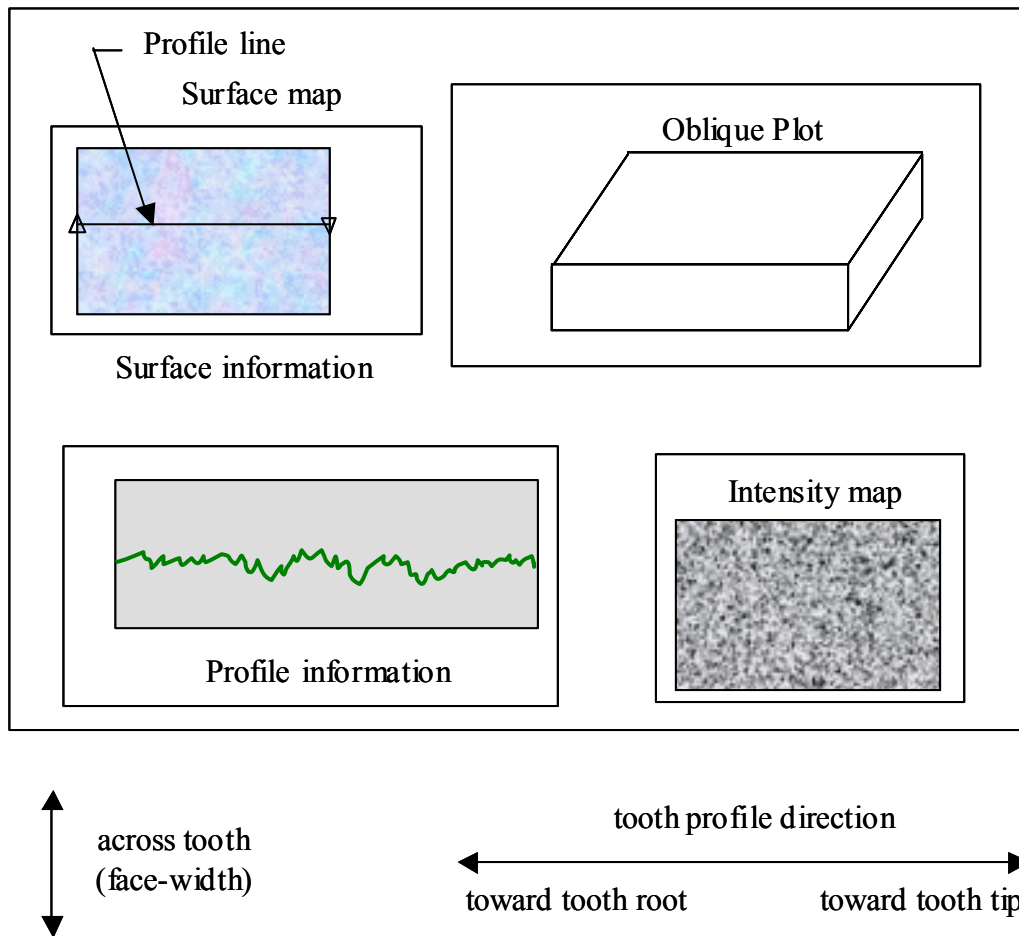


Figure 7.2.1.—Arrangement of the surface inspection data display and the orientation of the gear tooth relative to the measurement machine's coordinate system.

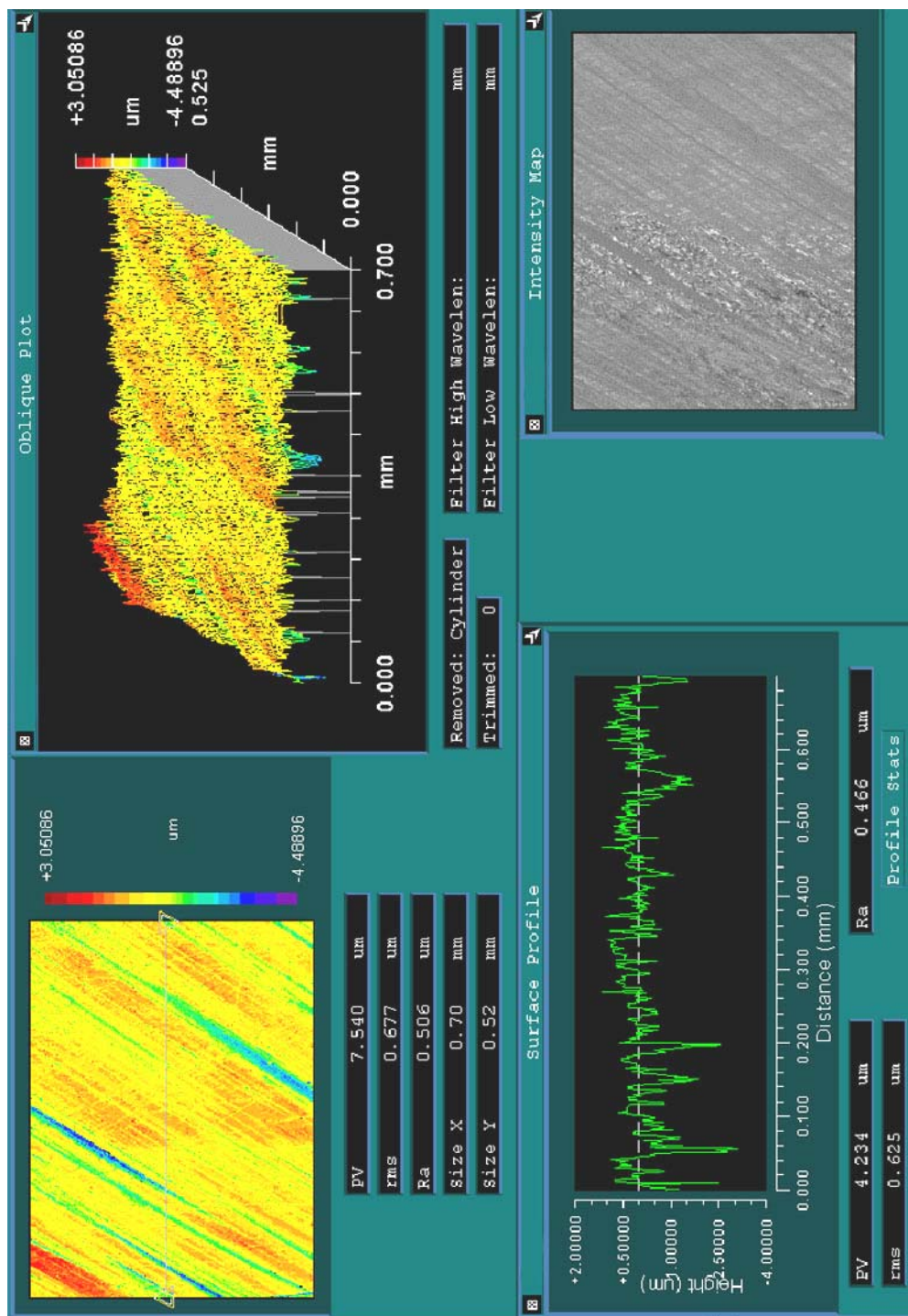


Figure 7.3.1.—Data display for the surface inspection of as-manufactured (un-run) area of ground

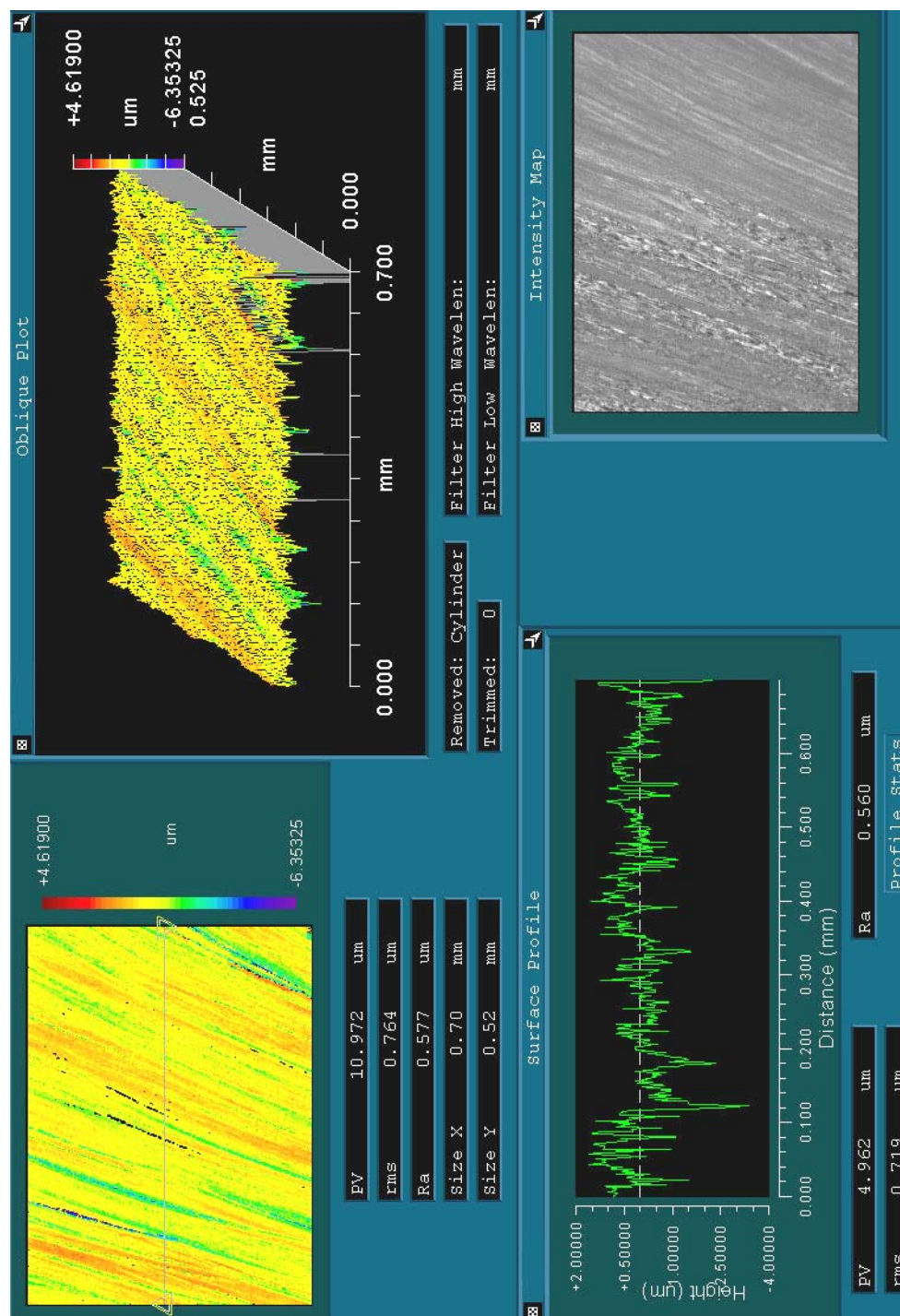


Figure 7.3.2.—Data display for the surface inspection of as-manufactured (un-run) area no. 1 of ground gear no. 2

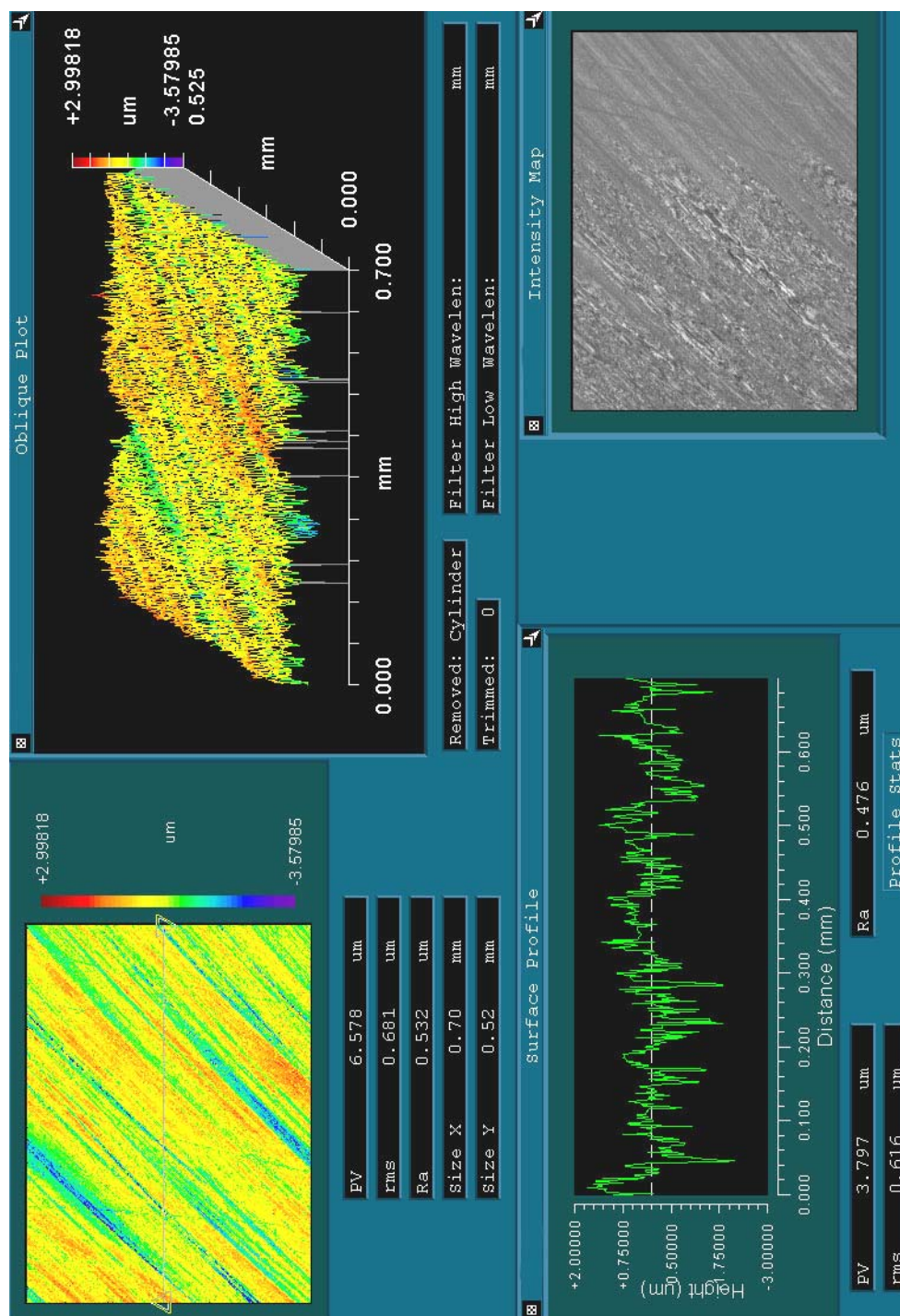


Figure 7.3.3.—Data display for the surface inspection of as-manufactured (un-run) area no. 2 for ground gear no. 2.

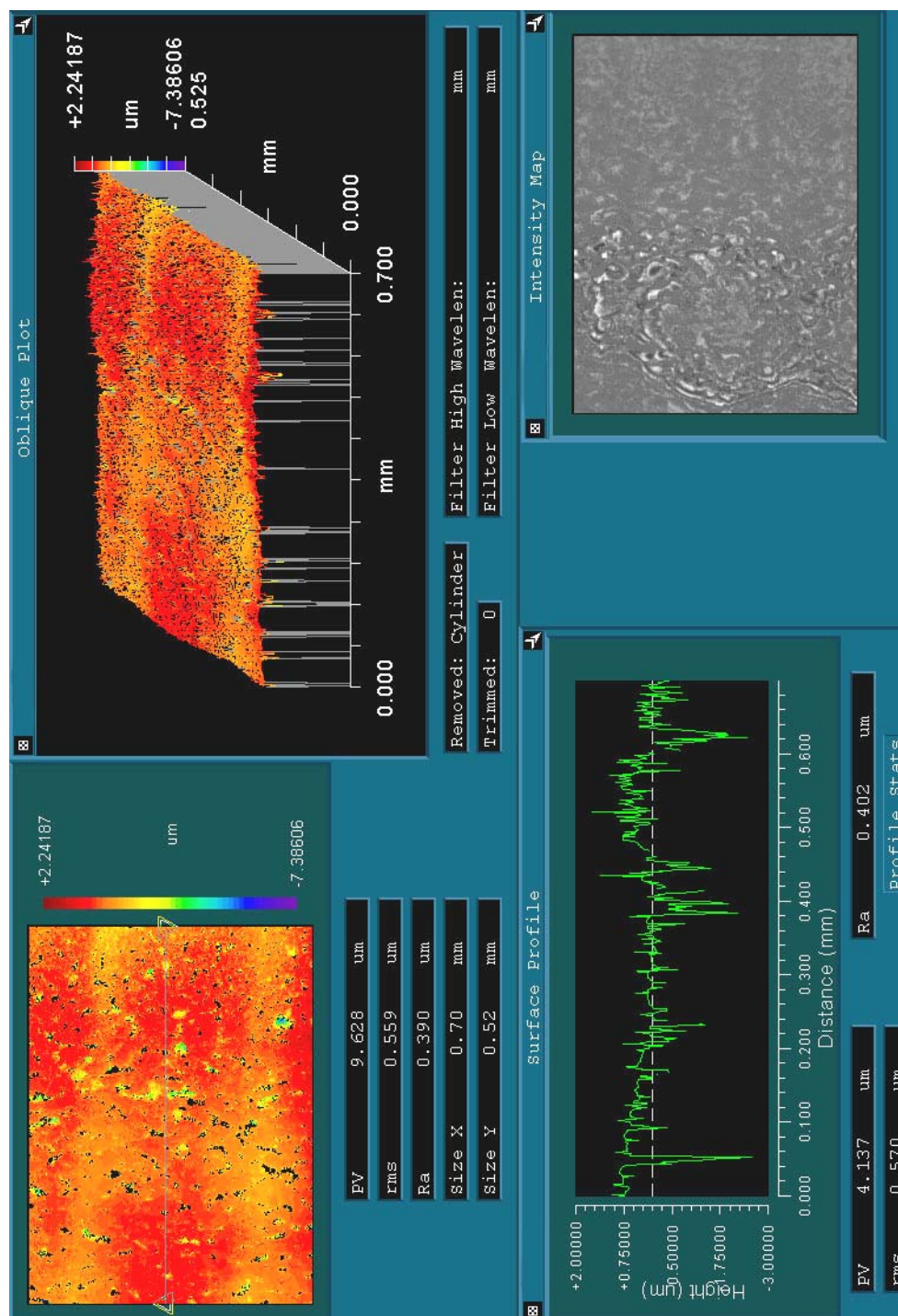


Figure 7.4.1.—Data display for the surface inspection of run-in ground gear, area no. 1.

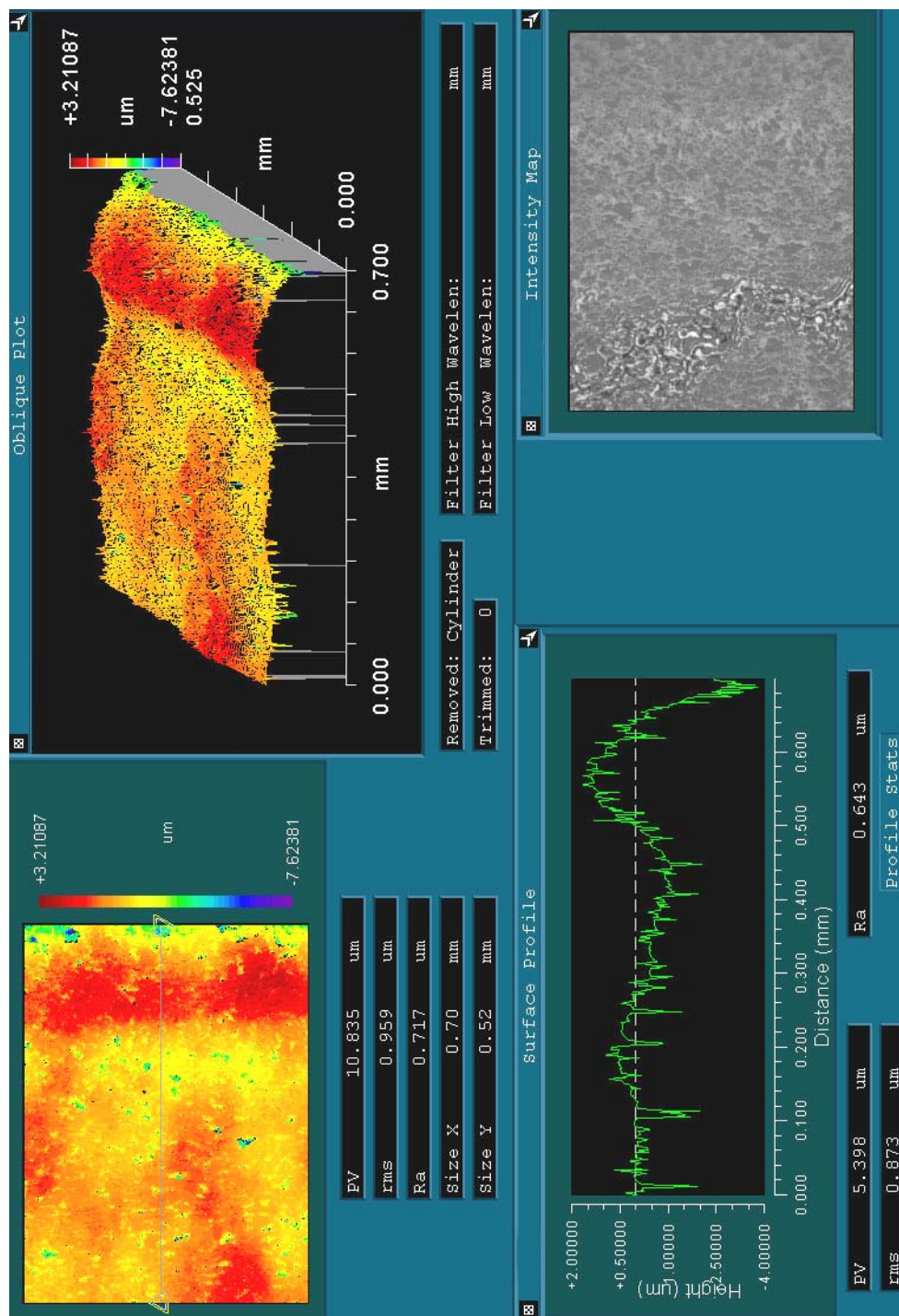


Figure 7.4.2.—Data display for the surface inspection of run-in area no. 2 for the ground gear.

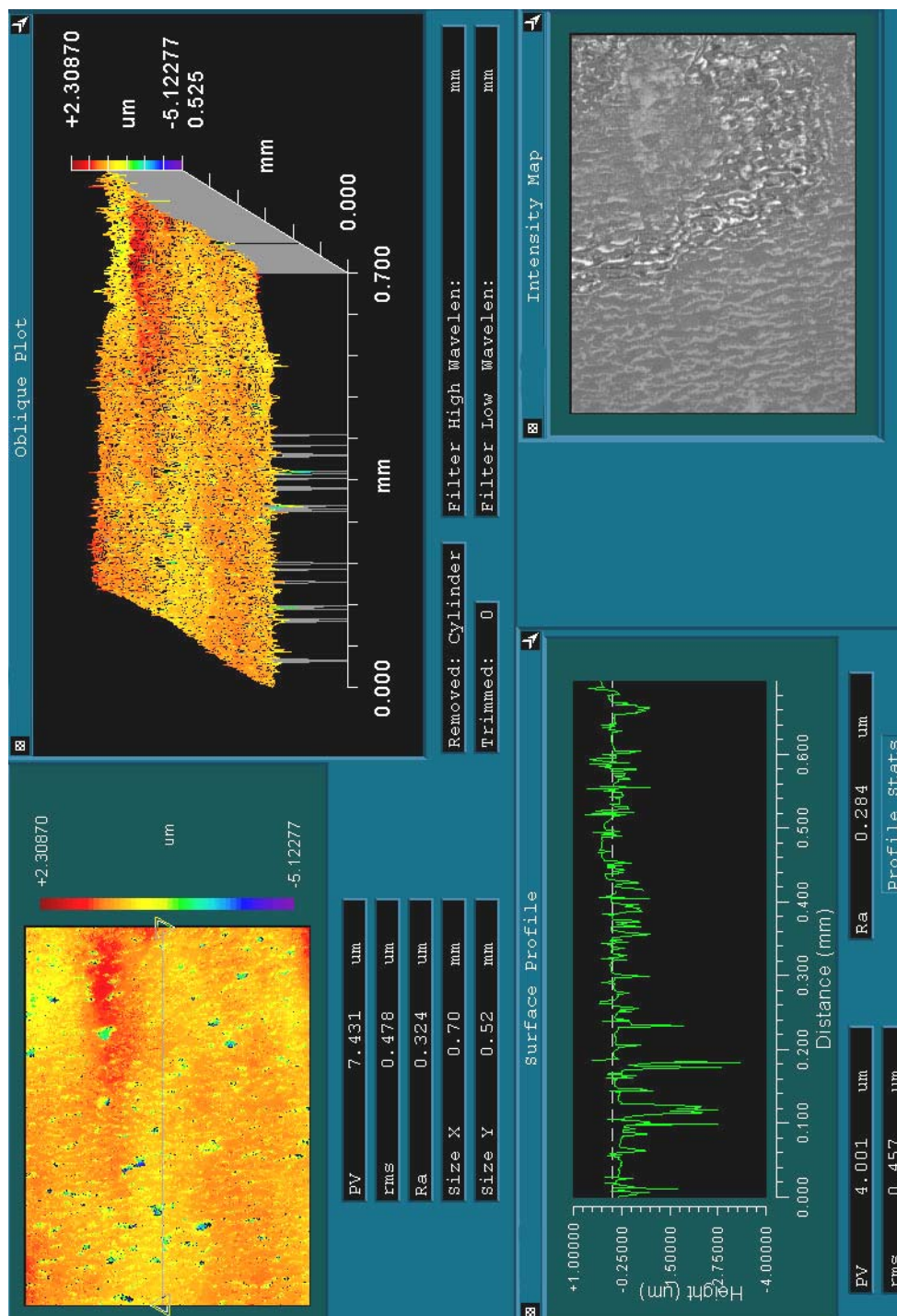


Figure 7.4.3.—Data display for the surface inspection of run-in area no. 3 for the ground gear.

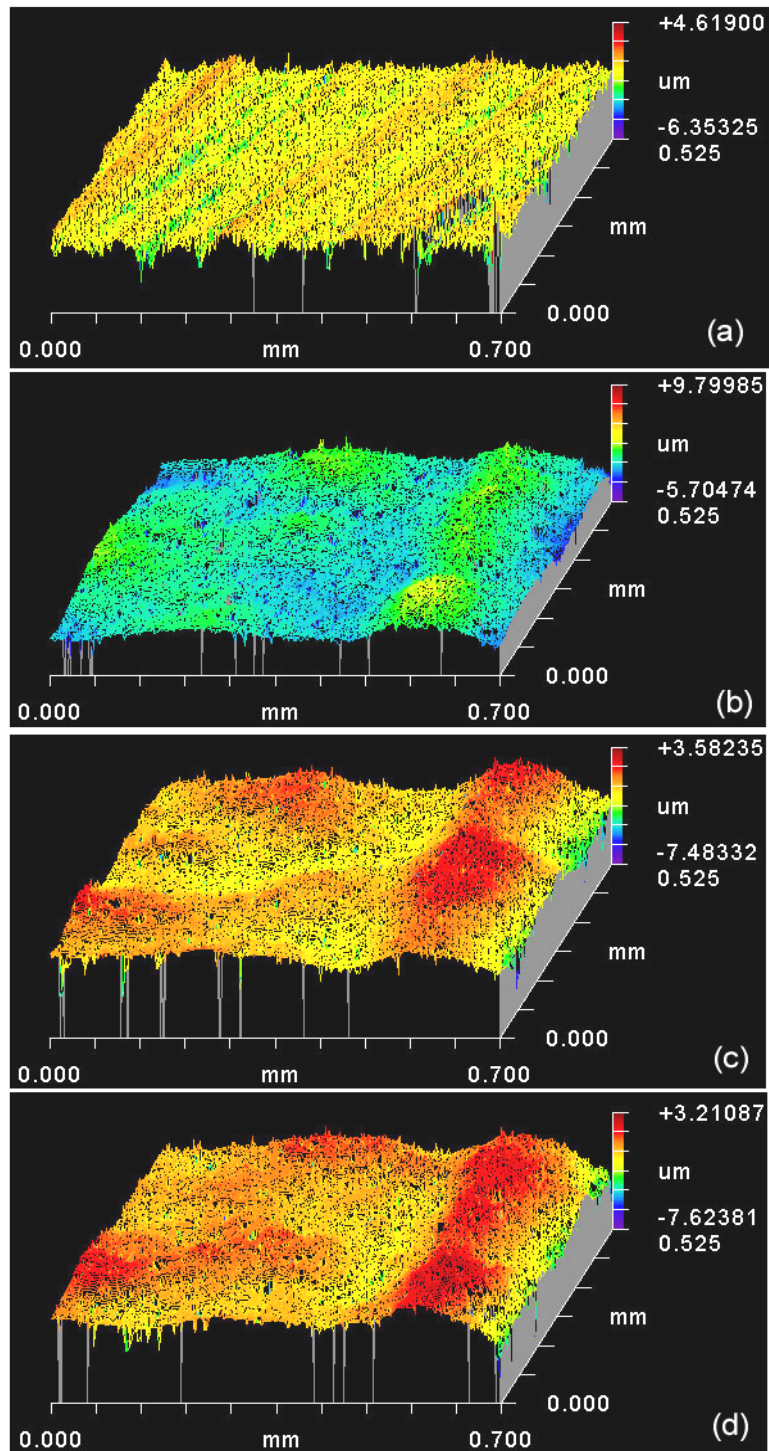


Figure 7.4.4.—Oblique plots of surface inspection data covering four separate areas across the face width near the pitch point along the involute profile for the ground gear.
 (a) Middle of the un-run portion of the tooth face.
 (b) Toward near edge of running track. (c) Middle of running track. (d) Toward far edge of running track.

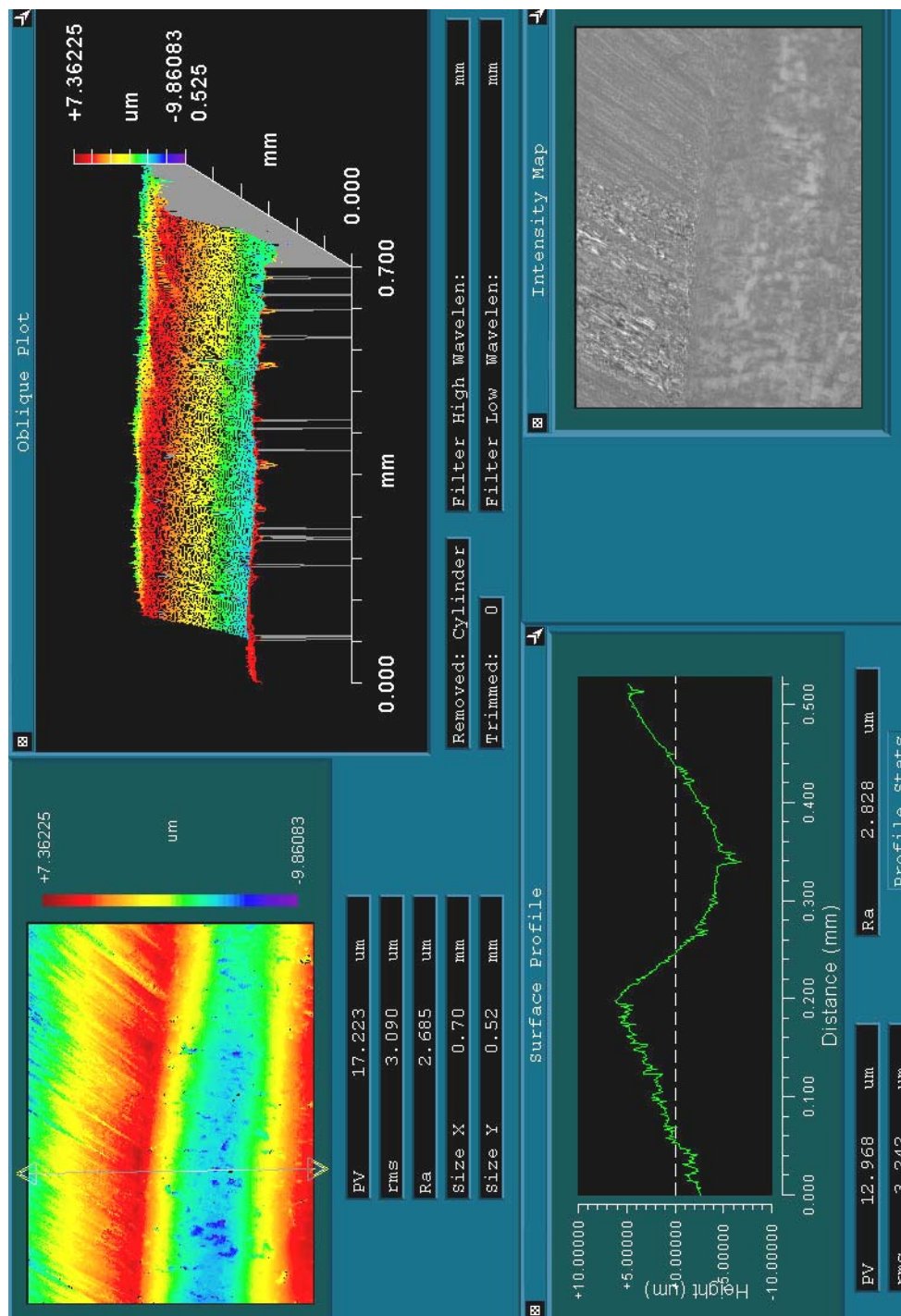


Figure 7.4.5.—Data display for the surface inspection of the ground gear at the area of transition from the run-in and un-run parts of the face width. Position along the involute profile is somewhat above the pitch point.

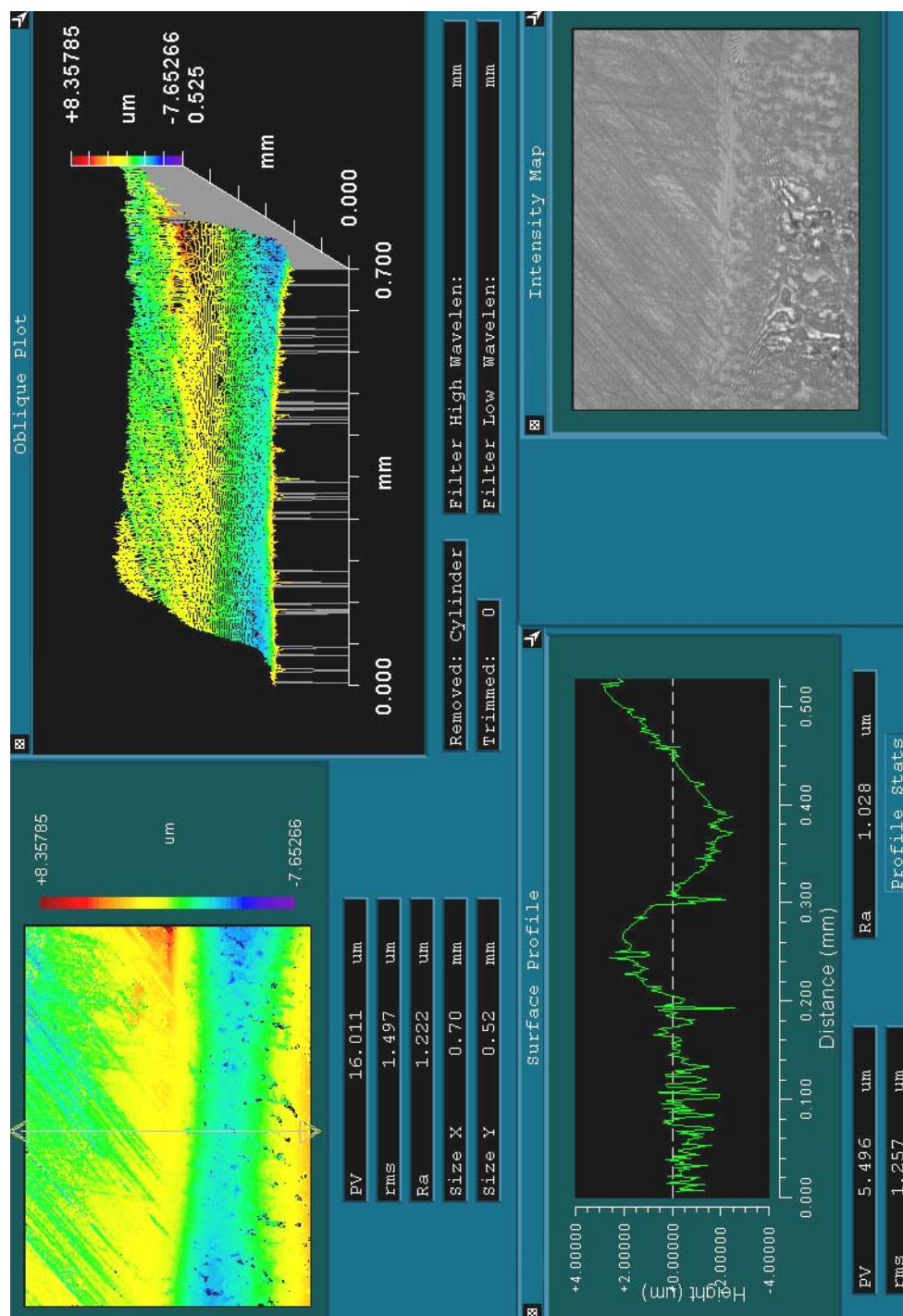


Figure 7.4.6.—Data display for the surface inspection of the ground gear at the area of transition from the run-in and un-run parts of the face width. Position along the involute profile is nearing the tooth tip.

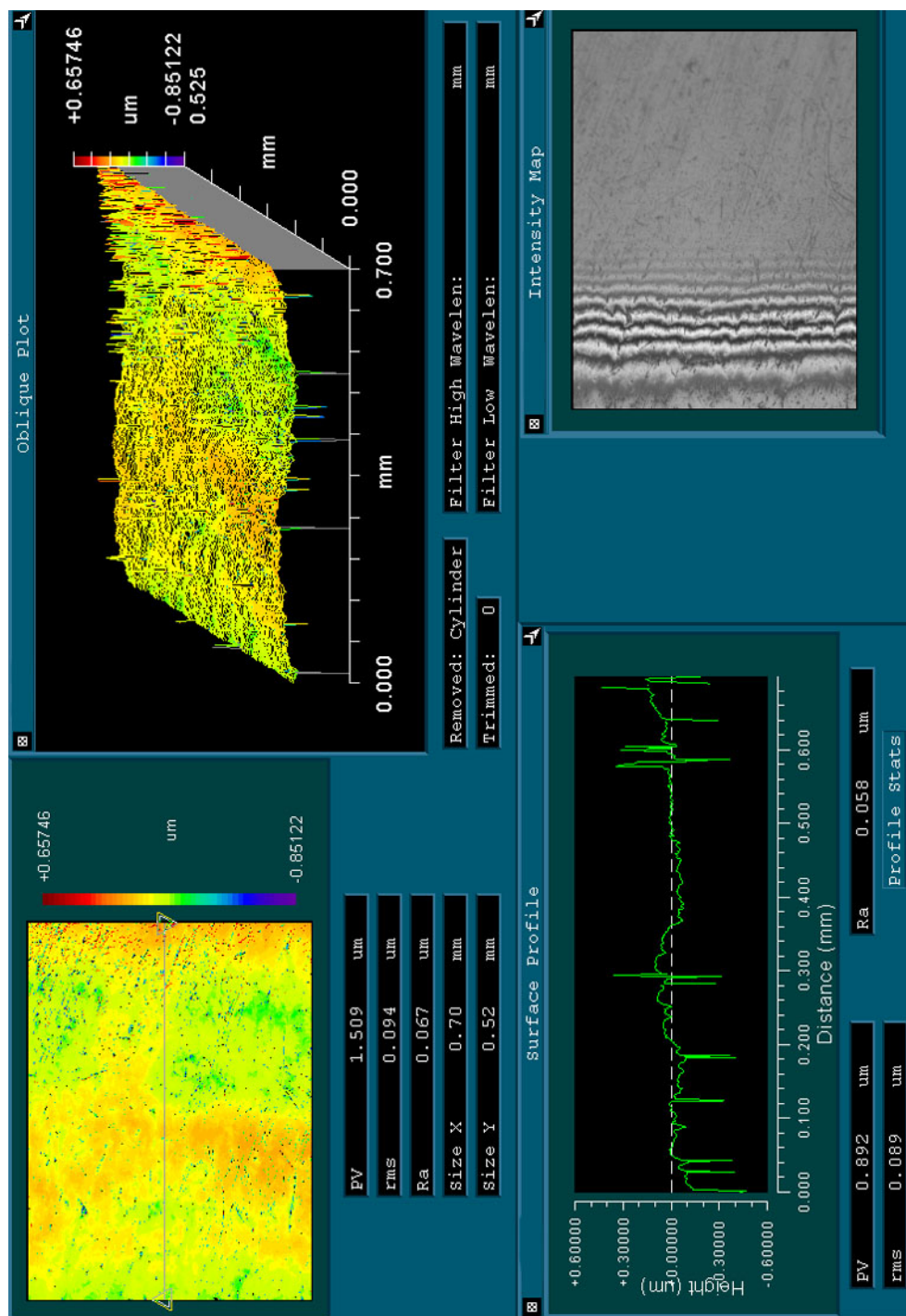


Figure 7.5.1.—Data display for the surface inspection of area no.1 of the superfinished gear.

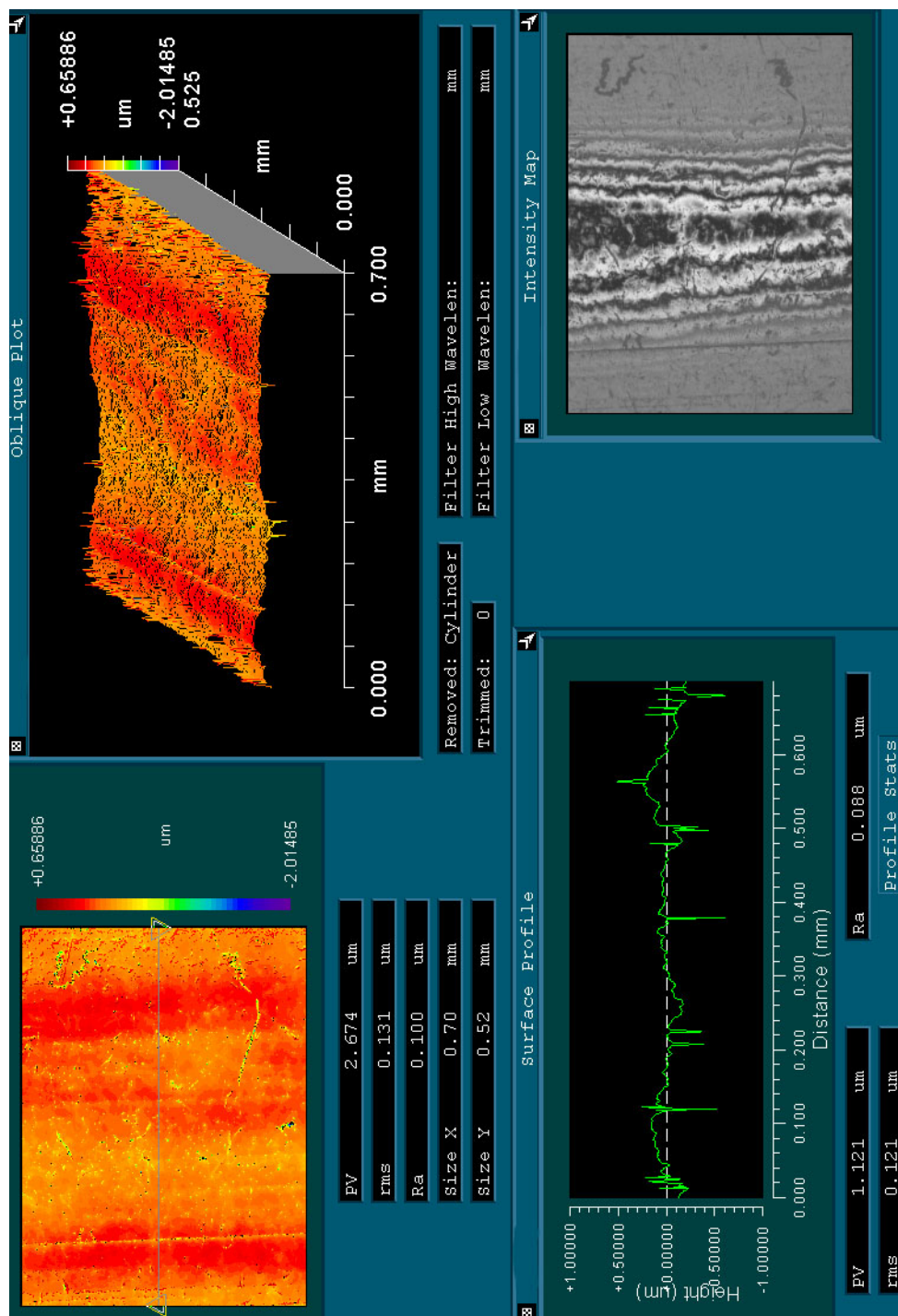


Figure 7.5.2.—Data display for the surface inspection of area no. 2 of the superfinished gear.

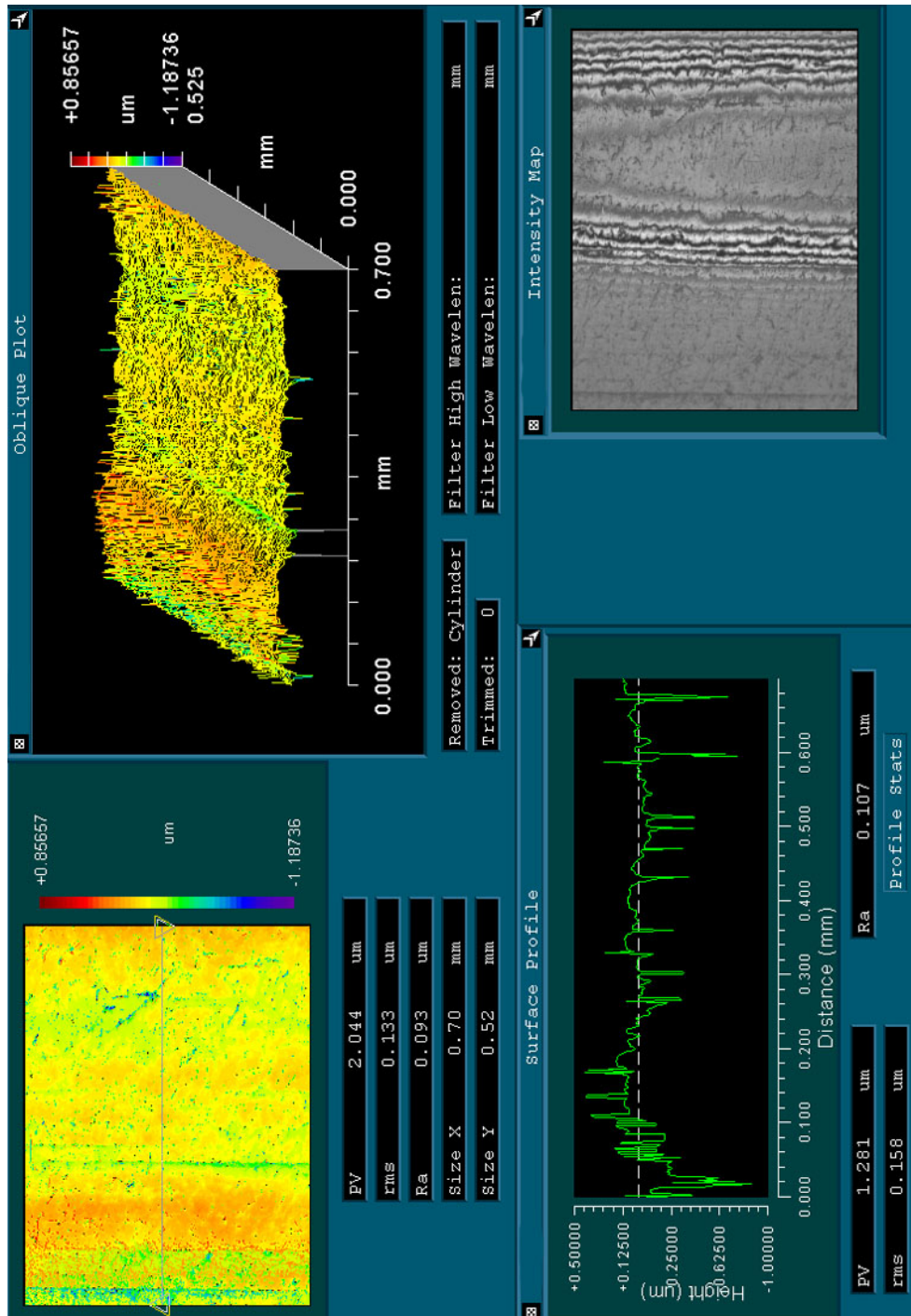


Figure 7.5.3.—Data display for the surface inspection of area no. 3 of the superfinished gear.

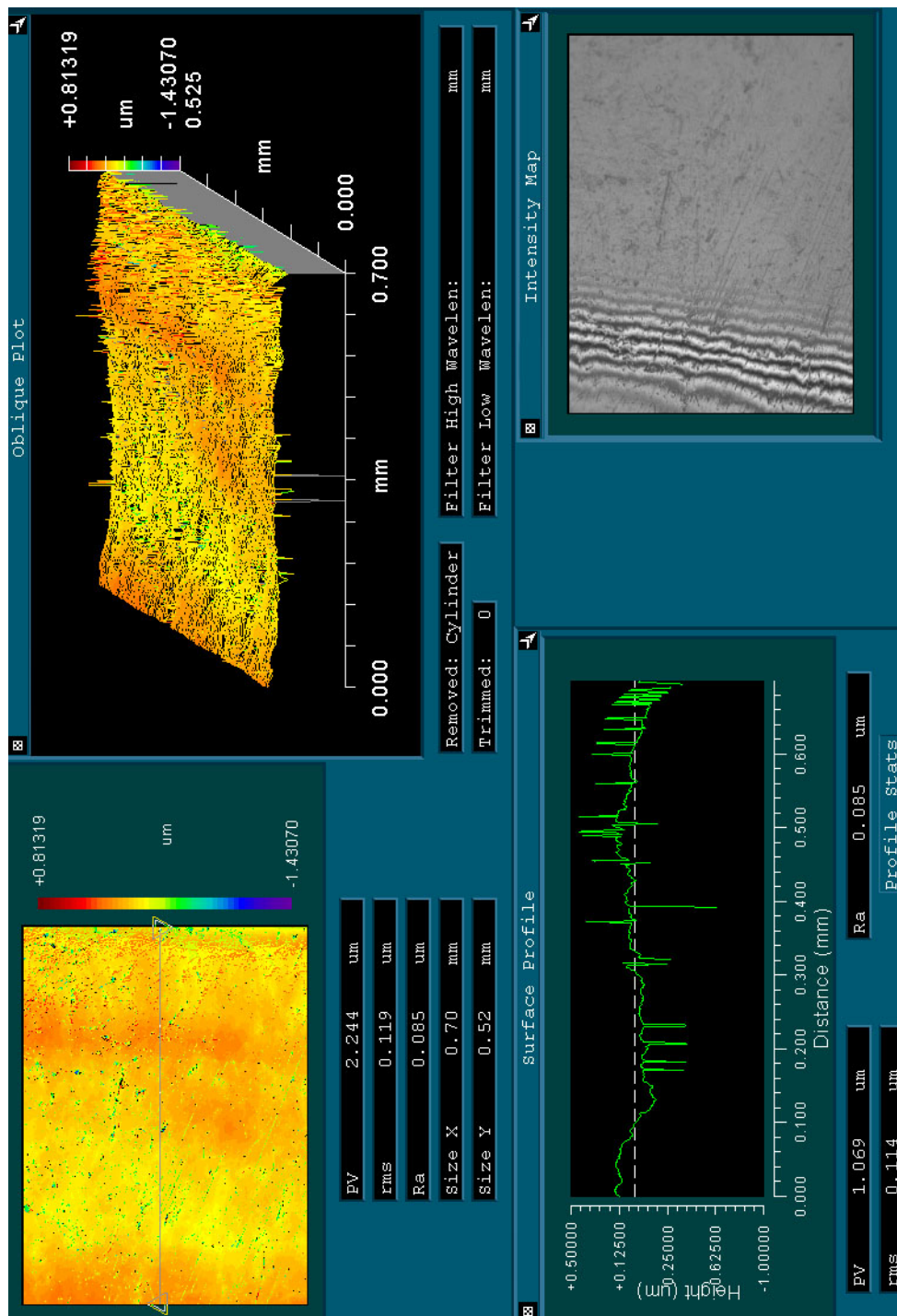


Figure 7.5.4.—Data display for the surface inspection of the superfinished gear. Position along the involute profile is approximately the pitch point and position across the face width is the mid-point of the face at the location of the un-run center track.

Chapter 8—Conclusions

The subject of this project was gear surface fatigue, with special attention given to the influence of surface roughness. Three groupings of complementary research were completed: (1) assessments and developments of statistical methods for gear fatigue data, (2) experimental evaluation of the fatigue lives of superfinished gears relative to ground gears, and (3) evaluation of the experimental condition by use of elasto-hydrodynamics, stress analysis, and surface topography inspections.

Consistent with other types of high cycle fatigue phenomena, the surface fatigue lives of nominally identical gears operated in a nominally identical fashion will vary greatly from one specimen to the next. Therefore, statistical concepts and methods are important to effectively evaluate and make use of data from surface fatigue experiments. In the present research, the Weibull distribution was selected as the distribution of choice for describing the distribution of gear fatigue lives. Evaluations and developments of statistical methods were completed with focus on two main goals. One goal was to estimate percentiles of the cumulative distribution function. The second goal was to compare two datasets and quantify, by a confidence level, the existence or absence of statistically significant differences in fatigue lives.

The evaluations of statistical methods were done making use of the Monte Carlo method to simulate random sampling. Sets of Monte Carlo simulations were completed to determine the sampling distributions for statistics of interest. The sampling distributions were evaluated for accuracy and precision of the sample statistics. The accuracy was evaluated using a newly introduced concept termed the mode-based bias. The method was applied to compare and contrast regression-based and likelihood-based methods for estimating parameters and determining confidence intervals. The usual assumption of a zero-valued Weibull location parameter was critically examined. To enable a method for the comparison of two datasets on the basis of a chosen percentile of interest, a new method was proposed, developed, and illustrated by example. The newly proposed method is an extension of likelihood-ratio based statistics. The significant contributions of the project to the state-of-the-art for statistical methods for gear fatigue data are as follows:

1. Software for the calculation of parameter estimates and confidence intervals were developed and validated.
2. The usual practice of describing gear fatigue data using the 2-parameter Weibull distribution rather than the more general 3-parameter Weibull distribution was critically examined. Although the experimental evidence and theoretical considerations both suggest that the true value of the threshold parameter is not identically zero, the assumption of a zero-valued threshold is a good approximation. It is recommended to model gear surface fatigue life distributions with a 2-parameter Weibull distribution and, thereby, as having a zero-valued threshold parameter.
3. Three methods for determining distribution parameters from sample data (two regression-based methods and a maximum likelihood based method) were evaluated for accuracy and precision. It is recommended to estimate the parameters of the Weibull distribution to describe gear surface fatigue life by using the maximum likelihood method.
4. A new method is proposed and developed for comparing two datasets. The purpose of the new method is to detect the existence of statistically significant differences in fatigue life properties. The method compares two datasets based on a selected quantile of the fatigue life cumulative distribution functions. A null hypothesis is set forth that a chosen quantile of the two populations are equal. By application of the method, one can either reject or fail-to-reject the null hypothesis on the basis of the experimental evidence as quantified by a confidence number.

The power density of a gearbox is an important consideration for many applications and is especially important for gearboxes used on aircraft. One factor that limits gearbox power density is the ability of the gear teeth to transmit power for the required number of cycles without pitting or spalling. Economical methods for improving surface fatigue lives of gears are therefore highly desirable. Works in the literature provide some evidence that the reduction of surface roughness improves the lubricating condition and

offers the possibility of increasing the surface fatigue lives of gears. However, there is little published data to quantify the improvement in life for case-carburized gears. Therefore, experiments were completed to quantify the surface fatigue lives of aerospace-quality gears that have been provided with an improved surface finish relative to conventionally ground gears.

A set of consumable-electrode vacuum-melted (CVM) AISI 9310 steel gears were ground and then provided with a near-mirror quality tooth surface by superfinishing. The gear teeth surface qualities were evaluated using metrology inspections, profilometry, and a mapping interferometric microscope. The gears were tested for surface fatigue in the NASA Glenn gear fatigue test apparatus at a load of 1.71 GPa (248 ksi) and at an operating speed of 10 000 rpm until failure or until survival of 300 million stress cycles. The lubricant used was a polyol-ester base stock meeting the specification DOD-L-85734. The failures were considered as failures of a two-gear system, and the data were fitted to a two-parameter Weibull distribution using the maximum likelihood method. The results of the present study were compared with the NASA Glenn gear fatigue data base. The following lists the specific results obtained for the experimental evaluations and the contributions of this project to the state-of-the-art.

1. The superfinishing treatment removed about 2 to 3 μm (79 to 118 $\mu\text{in.}$) of material from the tooth surfaces.
2. The superfinishing treatment reduced the mean roughness average (Ra) by a factor of about five and the mean 10-point parameter (Rz) value by a factor of about four.
3. The 10-percent life of the set of ground and superfinished gears of the present study was greater than the 10-percent life of the set of ground gears of the baseline study to a 90-percent confidence level.
4. In general, the life of the set of ground and superfinished gears of the present study was about five times greater than the life of the set of ground gears of the baseline study.
5. The set of ground and superfinished gears of the present study had lives greater than those of any other set of single-vacuum processed AISI 9310 gears tested to date using the NASA Glenn gear fatigue test apparatus.
6. The proportion of the gears operating for 300 million cycles without failure was considerably higher for the superfinished gears than was the proportion for any other set of ground AISI 9310 gears tested to date using the NASA Glenn gear fatigue test apparatus.
7. The lives of the CVM AISI 9310 ground and superfinished gears of the present study were of the order of magnitude of VIM-VAR AISI 9310 ground gears when tested using the NASA Glenn gear fatigue test apparatus.
8. There is strong evidence that superfinishing significantly improves the surface fatigue lives of case-carburized, ground, aerospace-quality AISI 9310 gears.

The results of the experimental research, as just stated, showed conclusively that gears with differing as-manufactured surface topographies can have dramatically differing performance characteristics. To apply these laboratory evaluations to products requires engineering understanding, analysis and judgment. Methodologies for evaluating the experimental condition were developed and applied so as to assist with the application of superfinishing for gears.

To establish the experimental conditions, the dynamic tooth forces were measured. Selected material properties were also measured. These measured data were then used as part of an analytical procedure to evaluate the stresses below the surface of the gear tooth at the selected critical position of the meshing cycle (the low-point of single tooth contact on the driving member). The analytical procedure included a consideration of the rough, lubricated surfaces. The result of the analysis procedure is a calculated stress tensor (included both stresses induced by load and residual stresses of the case-carburized tooth structure). Finally, a procedure is provided to determine critical-plane stress-based measures of load intensity as relates to fatigue damage.

The methodologies just described were applied to evaluate the experimental conditions used for the experimental research presented herein. First, the present research on the relation of surface roughness to

fatigue life was expressed in terms of specific film thickness. Data for gears were compared to that for bearings, and comments are provided concerning the limitations and use of the specific film concept. Secondly, a series of analysis were made using the simplifying assumption of smooth surfaces. It is demonstrated that such models do not adequately explain the experimentally observed improvements in fatigue life due to superfinishing of gears. Lastly, the stress condition is analyzed using a rough surface model. These analyses provide qualitative explanations for the improved fatigue lives. By evaluation of the experimental conditions, the following specific results and contributions of this project to the state-of-the-art were obtained.

1. Expressing available gear fatigue data in terms of specific film thickness, it is possible to draw a curve that is similar in shape to those that have been published for bearings, and the curve would fall within the confidence intervals for all the gear data.
2. Expressing available gear fatigue data in terms of specific film thickness, the quantitative life ratios for the superfinished and ground gears are similar to openly published data for roller bearing life ratios.
3. The present research shows that the magnitudes and locations (depths) of the critical stress indices are affected by residual stress fields. This is an improvement over previous methods that provided focus only to a preselected, single depth below the surface.
4. Calculations of the stress condition using a smooth surface model and both plane-strain and plane-stress responses to the applied loads showed differing characteristics. In general, the results from the plane-stress calculations tend to draw more attention to the condition at the surface, while for the plane-strain results one's attention may be drawn to the subsurface as the critical region with the highest indices of fatigue damage.
5. The selected multiaxial fatigue index used in this work offered little in the way of additional insight nor improved predictive capability.
6. Qualitative analysis using a smooth-surface model suggests that friction forces do not significantly influence the fatigue lives for well-lubricated contacts.
7. Quantitative predictions of increased fatigue life due to superfinishing are not adequately explained by analytical models that make use of smooth surface models even if the reduction of gear tooth friction due to superfinishing is accounted for in the calculations.
8. Elastohydrodynamic analyses using rough surfaces predict that even for near-mirror quality superfinished surfaces, the peak fluid pressures significantly exceed those for the ideally-smooth analysis. Peak pressure exceeded the smooth-surface maximum pressures by nearly 50 percent for the conditions analyzed herein.
9. Even if the undeformed roughness is less than the film thickness, still the roughness deforms significantly as the roughness passes into the contact.
10. The present research suggests that the superfinished gears had longer fatigue lives relative to ground gears largely due to a reduced rate of micropitting fatigue processes (surface or near-surface initiated fatigue) rather than due to any reduced rates of spalling (sub-surface) fatigue processes.

To help provide insight about the relationship of the surface topography and the resulting performance as was demonstrated in the present work, inspections were completed for both as-manufactured and run-in surfaces. The instrument used for the inspections employs scanning white light interferometry technology. The surface inspections showed that for the experimental conditions used in the present work for fatigue testing, the surface topography of the ground gears changed dramatically with running-in while the changes in surface topography for the superfinished gears was negligible. These inspections also provide some insight concerning the interactions of wear, plastic deformation, and fatigue phenomena. The inspections and discussions presented herein help to fully document the present work and, perhaps, will provide qualitative guidance for future theoretical and modeling treatments of gear contacts.

Bibliography

- Electronic Statistics Textbook*, Tulsa, Oklahoma: StatSoft, Inc., (1999). (Available online: <http://www.statsoft.com/textbook/stathome.html>) Accessed August 6, 2001.
- Life Adjustment Factors for Ball and Roller Bearings, An Engineering Design Guide*, ASME, New York, (1971).
- “NewView 5000 Operating Manual,” *OMP-0423G*, Zygo Corp., Middlefield, CT., (2001). Available online, http://support.zygo.com/home/site_dir.cgi, accessed April 24, 2002.
- “Strength and Hardness in Tool Steels,” *TECH-TOPICS Bulletin 112*, Timken Latrobe Steel Company, Latrobe, PA (not dated).
- Abernathy, R.; Breneman, J.; Medlin, C.; Reinman, G., *Weibull Analysis Handbook*, Air Force Wright Aeronautical Laboratories Report AFWAL-TR-83-2079, (1983). Available from National Technical Information Service, Washington, DC.
- Ai, X., “Random Surface Roughness on Fatigue Life of a Lubricated Contact,” *Trans. Of ASME, Journal of Tribology*, **120**, (1998).
- Alfredsson, B., “A Study on Contact Fatigue Mechanisms,” *Department of Solid Mechanics, Royal Institute of Technology, Stockholm, Sweden, Doctoral Thesis no. 44*, (2000).
- Allen, C., Townsend, D., Zaretsky, E., “Elastohydrodynamic Lubrication of a Spinning Ball in a Nonconforming Groove,” *ASME Journal of Lubrication Technology*, **92** [1], (1970).
- Anderson, N.; Lowenthal, S.; and Black, J., “An Analytical Method to Predict Efficiency of Aircraft Gearboxes,” *NASA TM-83716, AIAA Paper 84-1500*, (1984).
- Anderson, N.; Lowenthal, S., “Spur-Gear-System at Part and Full Load,” *NASA TP-1622*, (1980).
- Anderson, W.; Parker, R.; Zaretsky, E., “Effect of Nine Lubricants on Rolling-Contact Fatigue Life,” *NASA TN D-1404*, (1962).
- Anderson, W.; Sibley, L.; Zaretsky, E., “The Role of Elastohydrodynamic Lubrication in Rolling-Contact Fatigue,” *J. of Basic Eng.*, **85**, (1963).
- Anderson, W., “Elastohydrodynamic Lubrication Theory as a Design Parameter for Rolling Element Bearings,” *ASME Paper 70-DE-19, ASME Design Engineering Conference & Show, Chicago, IL*, ASME, New York, (1970).
- Anderson, W., et al., “Materials and Processing,” *STLE Life Factors For Rolling Bearings-STLE SP-34*, E. Zaretsky, ed., Society of Tribologists and Lubrication Engineers, Park Ridge, IL (1992).
- Batista, A., et al., “Contact Fatigue of Automotive Gears: Evolution and Effects of Residual Stresses Introduced by Surface Treatments,” *Fatigue Fract Engng Mater Struct*, **23**, (2000)
- Bergman, B., “Estimation of Weibull Parameters Using a Weight Function,” *Journal of Materials Science Letters*, **5** [6], (1986).
- Bingyuan, J., Xitong, Z., Lei, W., “Estimate by Fracture Mechanics Theory for Rolling Contact Fatigue Life and Strength of Case-Hardened Gear Materials With Computer,” *Engineering Fracture Mechanics*, **39** [5], (1991).
- Borsoff, V., “On the Mechanism of Gear Lubrication,” *J. of Basic Eng.*, (1959).
- Cann P., Ioannides E., Jacobson B., Lubrecht A., “The Lambda Ratio—A Critical Reexamination,” *Wear*, **175** [1-2], (1994)
- Cacciari, M.; Mazzanti, G.; Montanari, G., “Comparison of Maximum-Likelihood Unbiasing Methods for the Estimation of the Weibull Parameters,” *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996).
- Cacciari, M.; Montanari, G., “Discussion—Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and other Probability Paper,” *IEEE Trans. On Electrical Insulation*, **26** [6], (1991).
- Chiu, Y., “The Mechanism of Bearing Surface Fatigue—Experiments and Theories,” *Tribology Transactions*, **40** [4], (1997).

- Cioc, C., Moraru, L., Keith, T. and Kahraman, A., "An Integrated, Physics Based Lubrication Analysis Method for Helicopter Transmission Gear Surfaces with Asperities," *Proceedings of the American Helicopter Society 56th Annual Forum*, American Helicopter Society, (2000).
- Cioc, C., "A Deterministic Elastohydrodynamic Lubrication Model of High-Speed Transmission Components," private communications for the Rotorcraft Industry Technology Association, University of Toledo (2002).
- Cleveland, W., *Visualizing Data*. Hobart Press. New Jersey, (1993).
- Coddington, P., "Using Random Number Generators," (1995). (Available online <http://www.npac.syr.edu/projects/random/index.html>) Accessed Aug. 6, 2001.
- Coe, H., Zaretsky, E., "Effect Of Interference Fits On Roller Bearing Fatigue Life," *ASLE Transactions*, **30** [2], (1987)
- Cohen, A., "Maximum Likelihood Estimation in the Weibull Distribution Based On Complete and On Censored Samples," *Technometrics*, **7** [4], (1965).
- Cohen, A., "The Reflected Weibull Distribution," *Technometrics*, **15** [4], (1973).
- Coy, J.; Townsend, D.; Zaretsky, E., "Analysis of Dynamic Capacity of Low-Contact-Ratio Spur Gears Using Lundberg-Palmgren Theory," *NASA TN D-8029*, (1975).
- Coy, J., Townsend, D., Zaretsky, E., "Dynamic Capacity and Surface Fatigue Life for Spur and Helical Gears," *Journal of Lubrication Technology*, **98** [2], (1976).
- Deming, W., "On Probability as a Basis For Action," *The American Statistician*, 1975.
- Doganaksoy, N.; Schme, J., "Comparisons of Approximate Confidence Intervals for Distributions Used in Life-Data Analysis," *Technometrics*, **35** [2], (1993).
- Dowson, D., Higginson, G., *Elasto-Hydrodynamic Lubrication—SI Edition*, Pergamon Press, (1977).
- Efron, B.; Tibshirani, R., *An Introduction to the Bootstrap*, Chapman & Hall, New York (1993).
- Elcoate, C., Evans, H., Hughes, T., "Fully Coupled Elastohydrodynamic Solution Techniques for the Analysis of Real Rough Line Contacts Using Finite Element and Finite Difference Models," *Elastohydrodynamics '96—Proceedings of the 23rd Leeds-Lyon Symposium on Tribology*, D. Dowson, ed., Elsevier, (1997).
- Elcoate C., Evans H., Hughes T., and Snidle R., "Thin Film, Time Dependent, Micro-EHL Solutions With Real Surface Roughness" *Proceedings of the Leeds-Lyon Symposium, Lubrication at the Frontier, 1998*, Elsevier, (1999).
- Engelhardt, M.; Bain, L., "Simplified Statistical Procedures for the Weibull or Extreme-value Distribution," *Technometrics*, **19** [3], (1977).
- Engelhardt, M.; Bain, L., "On Prediction Limits for Samples From a Weibull or Extreme-value Distribution," *Technometrics*, **24** [2], (1982).
- Escobar, L., Meeker, W., "Statistical Prediction Based on Censored Life Data," *Technometrics*, **41** [2], (1999).
- Fan, H., Keer, L., Cheng, W., Cheng H., "Competition Between Fatigue Crack Propagation and Wear," *Journal of Tribology*, **115**, (1993).
- Fisher, R., "The General Sampling Distribution of the Multiple Correlation Coefficient," *Proceedings of the Royal Society of London, Ser. A*, **121**, (1928).
- Fishman, G.; Moore, L., "A Statistical Evaluation of Multiplicative Random Number Generators With Modulus 2311," *J. of the American Statistical Society*, **77**, (1982).
- Fishman, G.; Moore, L., "An Exhaustive Analysis of Multiplicative Congruential Random Number Generators with Modulus 231-1," *SIAM J. on Scientific and Statistical Computing*, **7**, (1986).
- Fothergill, J., "Estimating the Cumulative Probability of Failure Data Points to be Plotted on Weibull and other Probability Paper," *IEEE Trans. On Electrical Insulation*, **25** [3], (1990).
- Fushimi, M., "Random Number Generation with the Recursion $X_t = X_{t-3p} \oplus X_{t-3q}$," *J. of Computational and Applied Mathematics*, **31**, (1990).
- Graham, R., Olver, A., Macpherson, P., "An Investigation into the Mechanisms of Pitting in High-Hardness Carburized Steels," *ASME Paper 80-C2-DET*, ASME, New York, (1980).

- Handschuh, R., "Thermal Behavior of Spiral Bevel Gears," *Ph.D. Thesis, Case Western Reserve University, Cleveland, Ohio*, (1993).
- Hahn, G.; Meeker, W., *Statistical Intervals – A Guide for Practitioners*, John Wiley and Sons, New York, 1991.
- Hallinan, A., "A Review of the Weibull Distribution," *J. of Quality Technology*, **25** [2], (1993).
- Harter, H.; Moore, A., "Maximum-Likelihood Estimation of the Parameters of Gamma and Weibull Populations from Complete and from Censored Samples," *Technometrics*, **7** [4], (1965).
- Hartnett M., Kannel, J. "Contact Stresses Between Elastic Cylinders—A Comprehensive Theoretical And Experimental Approach," *Trans. of the ASME, Journal Of Lubrication Technology*, **103** [1], (1981).
- Hashimoto, M.; Hoyashita, S.; Iwata, J., "Studies on Improvement of Surface Durability of Case-Carburized Steel Gear—Effects of Surface Finish Processes upon Oil Film Formation," *AGMA Pub. 98FTM9*, (1998).
- Hertz, H., "On the Contact of Elastic Solids (in German)," *J. Reine und Angwandte Mathematik*, **92**, (1882).
- Hirose, H., "Maximum Likelihood Estimation in the 3-parameter Weibull Distribution," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996) and ERRATUM, **3** [2], (1996).
- Hirose, H., "Bias Correction for the Maximum Likelihood Estimates in the Two-parameter Weibull Distribution," *IEEE Trans. On Dielectrics and Electrical Insulation*, **6** [1], (1999).
- Hoyashita, S.; Hashimoto, M.; Seto, K., "Basic Studies on Fatigue Strength of Case-Hardened Gear Steel—Effects of Shot Peening and/or Barreling Processes," *AGMA Pub. 98FTM3*, (1998).
- Ishibashi, A.; Ezoe, S.; Tanaka, S., "Mirror Finishing of Tooth Surfaces Using a Trial Gear Grinder With Cubic-Born-Nitride Wheel," *ASME Pub. 84-DET-153*, ASME, New York, NY, 1984.
- Jacquelin, J., "A Reliable Algorithm for the Exact Median Rank Function," *IEEE Trans. On Electrical Insulation*, **28** [2], (1993).
- Jacquelin, J., "Inference of Sampling on Weibull Parameter Estimation," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [6], (1996).
- Jeng, S.; Meeker, W., "Comparisons of Approximate Confidence Interval Procedures for Type I Censored Data," *Technometrics*, **42** [2], (2000).
- Jones, D., Schott, G., Eds., *Miscellaneous Papers of H. Hertz*, Macmillan and Co., London, (1896).
- Johnson, K., *Contact Mechanics*, Cambridge University Press, (1985).
- Johnson, K., Shercliff, H., "Shakedown of 2-Dimensional Asperities in Sliding Contact," *International Journal of Mechanical Science*, **34** [5], 1992
- Johnson, L., *The Statistical Treatment of Fatigue Experiments*, Elsevier, New York, (1964).
- Kannel, J., Hartnett, M. "Theoretical And Experimental Evaluation Of Edge Stresses Under Severe Edge Loads," *ASLE Transactions*, **26** [1], (1983).
- Kannel, J., Tevaarwerk, J., "Subsurface Stress Evaluations Under Rolling/Sliding Contacts," *Trans. Of the ASME, J. of Tribology*, **106**, (1984).
- Kapoor, A., Johnson, K., "Effect of Changes in Contact Geometry on Shakedown of Surfaces in Rolling/Sliding Contact," *International Journal of Mechanical Science*, **34** [3], (1992).
- Kapoor, A., Morales-Espejel, G., Olver, A., "A Shakedown Analysis of Simple Spur Gears," *Tribology Transactions*, **45** [1], (2002).
- Keats, J.; Lawrence, F.; Wang, F., "Weibull Maximum Likelihood Parameter Estimates with Censored Data," *J. of Quality Technology*, **29** [1], (1997).
- Kotzalas, M., "A Theoretical Study of Residual Stress Effects on Fatigue Life Prediction," *Tribology Transactions*, **44** [4], (2001).
- Krantz, T., "Software for Statistical Analysis of Weibull Distributions With Application to Gear Fatigue Data—User's Manual with Verification," *NASA TM-2002-211109, ARL-TR-1338*, (2002).

- Kudish, I., Burris, K., "Modern State of Experimentation and Modeling in Contact Fatigue Phenomenon: Part I—Contact Fatigue vs. Normal and Tangential Contact and Residual Stresses, Nonmetallic Inclusions and Lubricant Contamination. Crack Initiation and Crack Propagation. Surface and Subsurface Cracks," *Tribology Transactions*, **42** [2], (2000).
- Kudish, I., Burris, K., "Modern State of Experimentation and Modeling in Contact Fatigue Phenomenon: Part II—Analysis of the Existing Statistical Mathematical Models of Bearing and Gear Fatigue Life. New Statistical Model of Contact Fatigue," *Tribology Transactions*, **43** [2], (2000).
- Kudish, I., "A New Statistical Model of Contact Fatigue," *Tribology Transactions*, **43** [4], (2000).
- Liu G., Wang, Q., Lin, C., "A Survey of Current Models for Simulating the Contact Between Rough Surfaces," *Tribology Transactions*, **42** [3], (1999).
- Liu, G., Wang, Q., "Thermoelastic Asperity Contacts, Frictional Shear, and Parameter Correlations," *Trans. Of the ASME, Journal of Tribology*, **122**, (2000).
- Liu, J.; Tallian, T.; McCool, J., "Dependence of Bearing Fatigue Life on Film Thickness to Surface Roughness Ratio," *ASLE Trans.*, **18** [2], (1975).
- Lubrecht, A., Venner, C., "Elastohydrodynamic Lubrication of Rough Surfaces," *Proc. Instn. Mech. Engrs., Part J*, **213**, (1999).
- Lundberg, G., Palmgren, A., "Dynamic Capacity of Rolling Bearings," *Ing. Vetenskaps Akad., Handl.*, **196**, (1947).
- Lundberg, G., Palmgren, A., "Dynamic Capacity of Rolling Bearings," *Ing. Vetenskaps Akad., Handl.*, **210**, (1952).
- McCool, J., "Evaluating Weibull Endurance Data by the Method of Maximum Likelihood," *ASLE Transactions*, **13** [3], (1970).
- Meeker, W.; Escobar, L., *Statistical Methods for Reliability Data*, John Wiley and Sons, New York, 1998.
- McCool, J., "Evaluating Weibull Endurance Data by the Method of Maximum Likelihood," *ASLE Transactions*, **13** [3], (1970).
- Montanari, G. et al., "In Search of Convenient Techniques for Reducing Bias in the Estimation of Weibull Parameters for Uncensored Tests," *IEEE Trans. On Dielectrics and Electrical Insulation*, **4** [3], (1997).
- Moyar, G., "A Critique of Rolling Contact Fatigue Criteria for Bearing Rings with Hoop Stresses," *Tribology Transactions*, **38** [2], (1995).
- Moyar, G., Stone, D., "An Analysis Of The Thermal Contributions To Railway Wheel Shelling," *Wear*, **144**, (1991).
- Moyer, C., "Using the Modified Lambda Ratio to Advance Bearing and Gear Performance," *SAE Technical Paper Series 901625, presented at the International Off-Highway & Powerplant Congress and Exposition*, SAE International, Warrendale, PA (1990).
- Moyer, C., Bahney, L., "Modifying the Lambda Ratio to Functional Line Contacts," *Tribology Transactions*, **33** [4], (1990).
- Nakatsuji, T.; Mori, A.; Shimotsuma, Y., "Pitting Durability of Electrolytically Polished Medium Carbon Steel Gears," *Tribology Trans.*, **38** [2], (1995) 223–232.
- Nakatsuji, T.; Mori, A., "Pitting Durability of Electrolytically Polished Medium Carbon Steel Gears—Succeeding Report," *Tribology Trans.*, **42** [2], (1999) 393–400.
- Nélias, D., et al., "Role of Inclusions, Surface Roughness and Operating Conditions on Rolling Contact Fatigue," *Trans. Of ASME, Journal of Tribology*, **121**, (1999).
- Nelson, W., "Theory and Applications of Hazard Plotting for Censored Failure Data," *Technometrics*, **14**, (1972).
- Nelson, W., *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*, John Wiley and Sons, New York, (1990).
- Olver, A., Spikes, H., Bower, A., and Johnson, K., "The Residual Stress Distribution in a Plastically Deformed Model Asperity," *Wear*, **107**, (1986).

- Patching M., Kweh C., Evans H., Snidle R., "Conditions For Scuffing Failure Of Ground And Superfinished Steel Disks At High Sliding Speeds Using A Gas-Turbine Engine Oil," *Trans. of the ASME, Journal Of Tribology*, **117** [3], (1995).
- Phansalkar, S., Kahrahan, A., Anderson, N., "A Comparative Study of EHL Behavior of Smooth and Asperity Contacts of Helical Gears and Rollers," 56th Annual Meeting of the Society of Tribologists and Lubrication Engineers, Orlando, Florida (2001).
- Pieracci, A., "Parameter Estimation for Weibull Probability Distribution Function of Initial Fatigue Quality," *AIAA Journal*, **33** [9], (1995).
- Present, D.; Newman, F.; Tyler, J.; Cuellar, J., "Advanced Chemical Characterization and Physical Properties of Eleven Lubricants," *NASA CR-168187*, (1983).
- Rebbechi, B., Oswald, F., Townsend, D., "Dynamic Measurements of Gear Tooth Friction and Load," *NASA-TM-103281, AVSCOM-TR-90-C-023*, (1990).
- Ringsberg, J. et al., "Prediction of Fatigue Crack Initiation For Rolling Contact Fatigue," *International Journal of Fatigue*, **22**, (2000).
- Ross, R., "Bias and Standard Deviation due to Weibull Parameter Estimation for Small Data Sets," *IEEE Trans. On Dielectrics and Electrical Insulation*, **3** [1], (1996).
- Sayles, R., "How Two- And Three-Dimensional Surface Metrology Data Are Being Used To Improve The Tribological Performance And Life Of Some Common Machine Elements," *Tribology International* **34** (5), (2001).
- Scibbe, W.; Townsend, D.; Aron, P., "Effect of Lubricant Extreme-Pressure Additives on Surface Fatigue Life of AISI 9310 Spur Gears," *NASA TP-2408*, (1984).
- Shimizu, S., "Fatigue Limit Concept and Life Prediction Model for Rolling Contact Machine Elements," *Tribology Transactions*, **45** [1], (2002).
- Skurka, J., "Elastohydrodynamic Lubrication of Roller Bearings," *Trans. Of the ASME, Journal of Lubrication Technology*, (1970).
- Smalley, A., et al., "Review of Failure Mechanisms in Highly Loaded Rolling and Sliding Contacts," *AD-657337*, Defense Technical Information Center, Ft. Belvoir, VA, (1967).
- Smith, J., Liu, C., "Stresses Due To Tangential and Normal Loads on an Elastic Solid With Applications to Some Contact Stress Problems," *Trans. ASME, Series E, J. of Applied Mechanics*, **20**, (1953).
- Snidle, R.; Evans, H.; Alanou, M., "The Effect of Superfinishing on Gear Tooth Profile," *AD-A327916*, Defense Technical Information Center, Ft. Belvoir, VA, (1997). Also available from NASA Scientific and Technical Information Center, <http://www.sti.nasa.gov/> accessed October 3, 2000.
- Snidle, R., Evans, H., Alanou, M., "Gears: Elastohydrodynamic Lubrication and Durability," *Proceedings of the Institution of Mechanical Engineers, Part C*, **214**, (2000).
- Socie, M., Marquis, G., *Multiaxial Fatigue*, Society of Automotive Engineers, Inc., Warrendale, Pa., (2000).
- Soman, K.; Misra, K., "A Least Square Estimation of Three Parameters of a Weibull Distribution," *Microelectronics Reliability*, **32** [3], (1992).
- Spikes, H., "Tribology Research in the Twenty-First Century," *Tribology International*, **34** (12), (2001).
- Tae, K., "Fatigue of Surface Engineered Steel in Rolling-Sliding Contact," *Ph.D. Thesis, Imperial College of Science, Technology, and Medicine, London, England*, (1999).
- Tallian, T., "Weibull Distribution of Rolling Contact Fatigue Life and Deviations There From," *ASLE Transactions*, **5**, (1962).
- Tallian, T., "On Competing Failure Modes in Rolling Contact," *ASLE Transactions*, **10** [4], (1967).
- Tallian, T., "The Influence of Asperity Statistics on Surface Distress and Spalling Life of Hertzian Contacts," *Tribology Transactions*, **36** [1], (1993).
- Tallian, T., "The Evolution of Rolling Bearing Life Models," *Society of Tribologists and Lubrication Engineers*, Park Ridge, IL, (2000).
- Tallian, T., "The Use of Microcontact Pressure Statistics in Fatigue Life Modeling of Rolling Contacts," *Tribology Transactions*, **44** [4], (2001).

- Tallian, T., Hoeprich, M., Kudish, I., "Discussions With Author's Closure—A New Statistical Model of Contact Fatigue," *Tribology Transactions*, **44** [4], (2001).
- Tanaka, S.; Ishibashi, A.; Ezoe, S., "Appreciable Increases in Surface Durability of Gear Pairs With Mirror-Like Finish," *ASME Pub. 84-DET-223*, ASME, New York, NY, (1984).
- Townsend, D.; Chevalier, J.; Zaretsky, E., "Pitting Fatigue Characteristics of AISI M-50 and Super Nitralloy Spur Gears," *NASA TN D-7261*, (1973).
- Townsend, D.; Bambergher, E.; Zaretsky, E., "Comparison of Pitting Fatigue life of Ausforged and Standard Forged AISI M-50 and AISI 9310 Spur Gears," *NASA TN-D-8030*, (1975).
- Townsend, D.; Coy, J.; Zaretsky, E., "Experimental and Analytical Load-Life Relation for AISI 9310 Steel Spur Gears," *NASA TM-X-73590*, (1977).
- Townsend, D., Coy, J., Zaretsky, E., "Experimental and Analytical Load-Life Relation for AISI 9310 Steel Spur Gears," *Trans. Of the ASME, Journal of Mechanical Design*, **100**, (1978).
- Townsend, D.; Parker, R.; Zaretsky, E., "Evaluation of CBS 600 Carburized Steel as a Gear Material," *NASA TP-1390*, (1979).
- Townsend, D.; Zaretsky, E., "Effect of Shot Peening on Surface Fatigue Life of Carburized and Hardened AISI 9310 Spur Gears," *NASA TP-2047*, (1982).
- Townsend, D.; Zaretsky, E., "Effect of Five Lubricants on Life of AISI 9310 Spur Gears," *NASA TP-2419*, (1985).
- Townsend, D., "Surface Fatigue Life and Failure Characteristics of EX-53, CBS 1000M, and AISI 9310 Gear Materials," *NASA TP-2513*, (1985).
- Townsend, D.; Bamberger, E., "Surface Fatigue Life of Carburized and Hardened M50NiL and AISI 9310 Spur Gears and Rolling-Contact Test Bars," *NASA TM-101979*, (1989).
- Townsend, D., "Surface Fatigue Life of Noninvolute, Low-Contact Ratio Gears," *NASA TM-103116*, (1990).
- Townsend, D.; Shimski, J., "Evaluation of Advanced Lubricants for Aircraft Applications Using Gear Surface Fatigue Tests," *NASA TM-104336*, (1991).
- Townsend, D., "Improvements in Surface Fatigue Life of Hardened Gears by High-Intensity Shot Peening," *NASA TM-105678*, (1992).
- Townsend, D.; Shimski, J., "Evaluation of the EHL Film Thickness and Extreme Pressure Additives on Gear Surface Fatigue Life," *NASA TM-106663*, (1994).
- Townsend, D., "Surface Fatigue Life of High Temperature Gear Materials," *NASA TM-106562, AIAA Paper 94-3078*, (1994).
- Townsend, D., Turza, A., and Chaplin, M., "The Surface Fatigue Life of Contour Induction Hardened AISI 1552 Gears," *NASA TM-107017, ARL-TR-808*, (1995).
- Way, S. "How To Reduce Surface Fatigue," *Machine Design*, (1939).
- Webster, M., Nobart, C., "An Experimentnal Investigation of Micropitting Using a Roller Disc Machine," *Tribology Transactions*, **38** [4], (1995).
- Weibull, W., "A Statistical Distribution of Wide Applicability," *J. Appl. Mech.*, **18**, (1951).
- Zaretsky, E., "Rolling Bearing and Gear Material," *Tribology for Aerospace Applications- STLE SP-37*, E. Zaretsky, ed., Society of Tribologists and Lubrication Engineers, Park Ridge, IL (1997).

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE October 2005		3. REPORT TYPE AND DATES COVERED Technical Memorandum
4. TITLE AND SUBTITLE The Influence of Roughness on Gear Surface Fatigue			5. FUNDING NUMBERS WBS-22-714-90-03 1L162211A47A	
6. AUTHOR(S) Timothy L. Krantz				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration John H. Glenn Research Center at Lewis Field Cleveland, Ohio 44135-3191			8. PERFORMING ORGANIZATION REPORT NUMBER E-15274	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001 and U.S. Army Research Laboratory Adelphi, Maryland 20783-1145			10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA TM-2005-213958 ARL-TR-3134	
11. SUPPLEMENTARY NOTES This report is an edited version of a document submitted as a dissertation in partial fulfillment of the requirements for the degree Doctor of Philosophy to Case Western Reserve University, Department of Mechanical and Aerospace Engineering, August 2002. Responsible person, Timothy L. Krantz, organization code RSM, 216-433-3580.				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category: 37 Available electronically at http://gltrs.grc.nasa.gov This publication is available from the NASA Center for AeroSpace Information, 301-621-0390.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Gear working surfaces are subjected to repeated rolling and sliding contacts, and often designs require loads sufficient to cause eventual fatigue of the surface. This research provides experimental data and analytical tools to further the understanding of the causal relationship of gear surface roughness to surface fatigue. The research included evaluations and developments of statistical tools for gear fatigue data, experimental evaluation of the surface fatigue lives of superfinished gears with a near-mirror quality, and evaluations of the experiments by analytical methods and surface inspections. Alternative statistical methods were evaluated using Monte Carlo studies leading to a final recommendation to describe gear fatigue data using a Weibull distribution, maximum likelihood estimates of shape and scale parameters, and a presumed zero-valued location parameter. A new method was developed for comparing two datasets by extending the current methods of likelihood-ratio based statistics. The surface fatigue lives of superfinished gears were evaluated by carefully controlled experiments, and it is shown conclusively that superfinishing of gears can provide for significantly greater lives relative to ground gears. The measured life improvement was approximately a factor of five. To assist with application of this finding to products, the experimental condition was evaluated. The fatigue life results were expressed in terms of specific film thickness and shown to be consistent with bearing data. Elastohydrodynamic and stress analyses were completed to relate the stress condition to fatigue. Smooth-surface models do not adequately explain the improved fatigue lives. Based on analyses using a rough surface model, it is concluded that the improved fatigue lives of superfinished gears is due to a reduced rate of near-surface micropitting fatigue processes, not due to any reduced rate of spalling (sub-surface) fatigue processes. To complete the evaluations, surface inspection were completed. The surface topographies of the ground gears changed substantially due to running, but the topographies of the superfinished gears were essentially unchanged with running.				
14. SUBJECT TERMS Fatigue tests; Gear; Weibull density functions; Stress analysis; Surface roughness			15. NUMBER OF PAGES 184	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT	

